BB

**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
A01N 43/04, A61K 31/70, C12Q 1/68

A1

(11) International Publication Number: **WO 98/18323**

(43) International Publication Date: 7 May 1998 (07.05.98)

(21) International Application Number: PCT/US97/19575

(22) International Filing Date: 28 October 1997 (28.10.97)

(30) Priority Data:

| | | |
|---|---|---|
| 08/739,150 | 28 October 1996 (28.10.96) | US |
| 08/759,739 | 6 December 1996 (06.12.96) | US |
| 08/891,928 | 14 July 1997 (14.07.97) | US |

(71) Applicant *(for all designated States except US)*: ASTRA AKTIEBOLAG [SE/SE]; S–151 85 Södertälje (SE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: SMITH, Douglas [US/US]; 2 Mayflower Lane, Gloucester, MA 01930 (US). ALM, Richard, A. [AU/US]; 28 Russet Hill Road, Ashland, MA 01721 (US).

(74) Agents: MANDRAGOURAS, Amy, E. et al.; Lahive & Cockfiel, LLP, 28 State Street, Boston, MA 02109 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

**Published**

*With international search report.*

*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *HELICOBACTER PYLORI* AND VACCINE COMPOSITIONS THEREOF

(57) Abstract

Recombinant or substantially pure preparations of *H. pylori* polypeptides are described. The nucleic acids encoding the polypeptides also are described. The *H. pylori* polypeptides are useful for diagnostics and vaccine compositions.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **AL** | Albania | **ES** | Spain | **LS** | Lesotho | **SI** | Slovenia |
| **AM** | Armenia | **FI** | Finland | **LT** | Lithuania | **SK** | Slovakia |
| **AT** | Austria | **FR** | France | **LU** | Luxembourg | **SN** | Senegal |
| **AU** | Australia | **GA** | Gabon | **LV** | Latvia | **SZ** | Swaziland |
| **AZ** | Azerbaijan | **GB** | United Kingdom | **MC** | Monaco | **TD** | Chad |
| **BA** | Bosnia and Herzegovina | **GE** | Georgia | **MD** | Republic of Moldova | **TG** | Togo |
| **BB** | Barbados | **GH** | Ghana | **MG** | Madagascar | **TJ** | Tajikistan |
| **BE** | Belgium | **GN** | Guinea | **MK** | The former Yugoslav Republic of Macedonia | **TM** | Turkmenistan |
| **BF** | Burkina Faso | **GR** | Greece | | | **TR** | Turkey |
| **BG** | Bulgaria | **HU** | Hungary | **ML** | Mali | **TT** | Trinidad and Tobago |
| **BJ** | Benin | **IE** | Ireland | **MN** | Mongolia | **UA** | Ukraine |
| **BR** | Brazil | **IL** | Israel | **MR** | Mauritania | **UG** | Uganda |
| **BY** | Belarus | **IS** | Iceland | **MW** | Malawi | **US** | United States of America |
| **CA** | Canada | **IT** | Italy | **MX** | Mexico | **UZ** | Uzbekistan |
| **CF** | Central African Republic | **JP** | Japan | **NE** | Niger | **VN** | Viet Nam |
| **CG** | Congo | **KE** | Kenya | **NL** | Netherlands | **YU** | Yugoslavia |
| **CH** | Switzerland | **KG** | Kyrgyzstan | **NO** | Norway | **ZW** | Zimbabwe |
| **CI** | Côte d'Ivoire | **KP** | Democratic People's Republic of Korea | **NZ** | New Zealand | | |
| **CM** | Cameroon | | | **PL** | Poland | | |
| **CN** | China | **KR** | Republic of Korea | **PT** | Portugal | | |
| **CU** | Cuba | **KZ** | Kazakstan | **RO** | Romania | | |
| **CZ** | Czech Republic | **LC** | Saint Lucia | **RU** | Russian Federation | | |
| **DE** | Germany | **LI** | Liechtenstein | **SD** | Sudan | | |
| **DK** | Denmark | **LK** | Sri Lanka | **SE** | Sweden | | |
| **EE** | Estonia | **LR** | Liberia | **SG** | Singapore | | |

# NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *HELICOBACTER PYLORI* AND VACCINE COMPOSITIONS THEREOF

## Background of the Invention

*Helicobacter pylori* is a gram-negative, S-shaped, microaerophilic bacterium that was discovered and cultured from a human gastric biopsy specimen. (Warren, J.R. and B. Marshall, (1983) *Lancet* 1: 1273-1275; and Marshall et al., (1984) *Microbios Lett.* 25: 83-88). *H. pylori* has been strongly linked to chronic gastritis and duodenal ulcer disease. (Rathbone et. al., (1986) *Gut* 27: 635-641). Moreover, evidence is accumulating for an etiologic role of *H. pylori* in nonulcer dyspepsia, gastric ulcer disease, and gastric adenocarcinoma. (Blaser M. J., (1993) *Trends Microbiol.* 1: 255-260). Transmission of the bacteria occurs via the oral route, and the risk of infection increases with age. (Taylor, D.N. and M. J. Blaser, (1991) *Epidemiol. Rev* 13: 42-50). *H. pylori* colonizes the human gastric mucosa, establishing an infection that usually persists for decades. Infection by *H. pylori* is prevalent worldwide. Developed countries have infection rates over 50% of the adult population, while developing countries have infection rates reaching 90% of the adults over the age of 20. (Hopkins R. J. and J. G. Morris (1994) *Am. J. Med.* 97: 265-277).

The bacterial factors necessary for colonization of the gastric environment, and for virulence of this pathogen, are poorly understood. Examples of the putative virulence factors include the following: urease, an enzyme that may play a role in neutralizing gastric acid pH (Eaton et al., (1991) *Infect. Immunol.* 59: 2470-2475; Ferrero, R.L. and A. Lee (1991) *Microb. Ecol. Hlth. Dis.* 4: 121-134; Labigne et al., (1991) *J. Bacteriol.* 173: 1920-1931); the bacterial flagellar proteins responsible for motility across the mucous layer. (Hazell et al., (1986) *J. Inf. Dis.* 153: 658-663; Leying et al., (1992) *Mol. Microbiol.* 6: 2863-2874; and Haas et al., (1993) *Mol. Microbiol.* 8: 753-760); Vac A, a bacterial toxin that induces the formation of intracellular vacuoles in epithelial cells (Schmitt, W. and R. Haas, (1994) *Molecular Microbiol.* 12(2): 307-319); and several gastric tissue-specific adhesins. (Boren et al., (1993) *Science* 262: 1892-1895; Evans et al., (1993) *J. Bacteriol.* 175: 674-683; and Falk et al., (1993) *Proc. Natl. Acad. Sci.* USA 90: 2035-203).

Numerous therapeutic agents are currently available that eradicate *H. pylori* infections *in vitro*. (Huesca et. al., (1993) *Zbl. Bakt.* 280: 244-252; Hopkins, R. J. and J. G. Morris, supra). However, many of these treatments are suboptimally effective *in vivo* because of bacterial resistance, altered drug distribution, patient non-compliance or poor drug availabilty. (Hopkins, R. J. and J. G. Morris, supra). Treatment with antibiotics combined with bismuth are part of the standard regime used to treat *H. pylori* infection.

(Malfertheiner, P. and J. E. Dominguez-Munoz (1993) *Clinical Therapeutics* 15 Supp. B: 37-48). Recently, combinations of a proton pump inhibitors and a single antibiotic have been shown to ameliorate duodenal ulcer disease. (Malfertheiner, P. and J. E. Dominguez-Munoz supra). However, methods employing antibiotic agents can have the problem of the emergence of bacterial strains which are resistant to these agents. (Hopkins, R. J. and J. G. Morris, supra). These limitations demonstrate that new more effective methods are needed to combat *H. pylori* infections *in vivo*. In particular, the design of new vaccines that may prevent infection by this bacterium is highly desirable.

## Summary of the Invention

This invention relates to novel genes, e.g., genes encoding polypeptides such as bacterial surface proteins, from the organism *Helicobacter pylori* (*H. pylori*), and other related genes, their products, and uses thereof. The nucleic acids and peptides of the present invention have utility for diagnostic and therapeutics for *H. pylori* and other *Helicobacter* species. They can also be used to detect the presence of *H. pylori* and other *Helicobacter* species in a sample; and for use in screening compounds for the ability to interfere with the *H. pylori* life cycle or to inhibit *H. pylori* infection. More specifically, this invention features compositions of nucleic acids corresponding to entire coding sequences of *H. pylori* proteins, including surface or secreted proteins or parts thereof, nucleic acids capable of binding mRNA from *H. pylori* proteins to block protein translation, and methods for producing *H. pylori* proteins or parts thereof using peptide synthesis and recombinant DNA techniques. This invention also features antibodies and nucleic acids useful as probes to detect *H. pylori* infection. In addition, vaccine compositions and methods for the protection or treatment of infection by *H. pylori* are within the scope of this invention.

## Detailed Description of the Drawings

Figure 1 is a bar graph that depicts the antibody titer in serum of mice following immunization with specific *H. pylori* antigens.

Figure 2 is a bar graph that depicts the antibody titer in mucous of mice following immunization with specific *H. pylori* antigens.

Figure 3 is a bar graph that depicts therapeutic immunization of *H. pylori* infected mice with specific antigens dissolved in HEPES buffer.

Figure 4 is a bar graph that depicts therapeutic immunization of *H. pylori* infected mice with specific antigens dissolved in buffer containing DOC.

Figure 5 depicts the amino acid sequence alignment in a portion of the sequence of five *H. pylori* proteins (depicted in the single letter amino acid code; shown N-terminal to C-terminal, left to right).

Figure 6 depicts the amino acid sequence alignment in a portion of the sequence of four *H. pylori* proteins (depicted in the single letter amino acid code; shown N-terminal to C-terminal, left to right).

Figure 7 depicts the amino acid sequence alignment in a portion of the sequence of two *H. pylori* proteins (depicted in the single letter amino acid code; shown N-terminal to C-terminal, left to right).

Figure 8 depicts the amino acid sequence alignment in a portion of the sequence of two *H. pylori* proteins (depicted in the single letter amino acid code; shown N-terminal to C-terminal, left to right).

Detailed Description of the Invention

In one aspect, the invention features a recombinant or substantially pure preparation of *H. pylori* polypeptide of SEQ ID NO: 74. The invention also includes substantially pure nucleic acid encoding an *H. pylori* polypeptide of SEQ ID NO: 74, such nucleic acid is contained in SEQ ID NO: 1. The *H. pylori* polypeptide sequences of the invention described herein are contained in the Sequence Listing, and the nucleic acids encoding *H. pylori* polypeptides of the invention are contained in the Sequence Listing.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 75, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 2.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 76, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 3.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 77, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 4.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 78, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 5.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 79, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 6.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 80, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 7.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 81, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 8.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 82, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 9.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 83, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 10.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 84, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 11.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 85, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 12.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 86, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 13.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 87, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 14.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 88, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 15.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 89, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 16.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 90, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 17.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 91, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 18.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 92, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 19.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 93, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 20.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 94, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 21.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 95, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 22.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 96, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 23.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 97, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 24.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 98, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 25.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 99, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 26.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 100, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 27.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 101, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 28.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 102, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 29.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 103, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 30.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 104, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 31.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 105, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 32.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 106, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 33.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 107, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 34.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 108, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 35.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 109, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 36.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 110, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 37.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 111, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 38.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 112, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 39.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 113, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 40.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 114, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 41.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 115, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 42.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 116, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:43.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 117, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 44.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 118, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 45.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 119, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 46.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 120, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 47.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 121, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 48.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 122, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 49.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 123, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 50.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 124, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 51.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 125, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 52.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 126, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 53.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 127, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 54.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 128, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 55.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 129, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 56.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 130, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 57.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 131, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 58.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 132, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 59.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 133, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 60.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 134, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 61.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 135, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 62.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 136, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 63.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 137, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 64.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 138, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 65.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 139, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 66.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 140, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 67.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 141, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 68.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 142, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 69.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 143, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 70.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 144, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 71.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 145, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 72.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 146, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 73.

Particularly perferred is an isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cell envelope polypeptide or a fragment thereof. Such nucleic acid is selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 25, SEQ ID NO: 48, SEQ ID NO: 16, SEQ ID NO: 10, SEQ ID NO: 45, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 7, SEQ ID NO: 39, SEQ ID NO: 55, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 1, SEQ ID NO: 42, SEQ ID NO: 14, SEQ ID NO: 43, SEQ ID NO: 11, SEQ ID NO: 71, SEQ ID NO: 17, SEQ ID NO: 57, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 8, and SEQ ID NO: 21.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof encoded by the nucleic acid selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 25, and SEQ ID NO: 48.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof encoded by the nucleic acid selected from the group consisting of SEQ ID NO: 16, SEQ ID NO: 10,

SEQ ID NO: 45, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 7, SEQ ID NO: 39, SEQ ID NO: 55, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 1, SEQ ID NO: 42, SEQ ID NO: 14, SEQ ID NO: 43, SEQ ID NO: 11, and SEQ ID NO: 71.

In another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof encoded by the nucleic acid selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 42, SEQ ID NO: 14, SEQ ID NO: 43, SEQ ID NO: 11, and SEQ ID NO:71.

In yet another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof encoded by the nucleic acid selected from the group consisting of SEQ ID NO: 16, SEQ ID NO: 45, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 7, SEQ ID NO: 39, SEQ ID NO: 55, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, and SEQ ID NO: 58.

Particularly preferred is an isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* secreted polypeptide or a fragment thereof. Such nucleic acid is selected from the group consisting of SEQ ID NO: 72, SEQ ID NO: 32, SEQ ID NO: 51, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 22, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40. SEQ ID NO: 41, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 49, SEQ ID NO: 53, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67, and SEQ ID NO: 68.

Particularly preferred is an isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cellular polypeptide or a fragment thereof. Such nucleic acid is selected from the group consisting of SEQ ID NO: 12, SEQ ID NO: 15, SEQ ID NO: 20, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 47, SEQ ID NO: 50, SEQ ID NO: 60, SEQ ID NO: 64, SEQ ID NO: 69, SEQ ID NO: 70, and SEQ ID NO: 73.

Particularly preferred is a purified or isolated *H. pylori* cell envelope polypeptide or a fragment thereof, wherein the polypeptide is selected from the group consisting of SEQ ID NO: 76, SEQ ID NO: 98, SEQ ID NO: 121, SEQ ID NO: 89, SEQ ID NO: 83, SEQ ID NO: 118, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 80, SEQ ID NO: 112, SEQ ID NO: 128, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 101, SEQ ID

NO: 103, SEQ ID NO: 125, SEQ ID NO: 127, SEQ ID NO: 129, SEQ ID NO: 131, SEQ ID NO: 74, SEQ ID NO: 115, SEQ ID NO: 87, SEQ ID NO: 116, SEQ ID NO: 84, SEQ ID NO: 144, SEQ ID NO: 90, SEQ ID NO: 130, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 81, and SEQ ID NO: 94.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 76, SEQ ID NO: 98, and SEQ ID NO: 121.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 89, SEQ ID NO: 83, SEQ ID NO: 118, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 80, SEQ ID NO: 112, SEQ ID NO: 128, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 125, SEQ ID NO: 127, SEQ ID NO: 129, SEQ ID NO: 131, SEQ ID NO: 74, SEQ ID NO: 115, SEQ ID NO: 87, SEQ ID NO: 116, SEQ ID NO: 84, SEQ ID NO: 144, SEQ ID NO: 90, and SEQ ID NO: 130.

In another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof selected from the group consisting of SEQ ID NO: 74, SEQ ID NO: 115, SEQ ID NO: 87, SEQ ID NO: 116, SEQ ID NO: 84, and SEQ ID NO:144.

In another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof selected from the group consisting of SEQ ID NO: 89, SEQ ID NO: 118, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 80, SEQ ID NO: 112, SEQ ID NO: 128, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 125, SEQ ID NO: 127, SEQ ID NO: 129, SEQ ID NO: 131.

Particularly preferred is a purified or isolated *H. pylori* secreted polypeptide or a fragment thereof, wherein the polypeptide is selected from the group consisting of SEQ ID NO: 145, SEQ ID NO: 105, SEQ ID NO: 124, SEQ ID NO: 75, SEQ ID NO: 77, SEQ ID NO: 82, SEQ ID NO: 86, SEQ ID NO: 95, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 107, SEQ ID NO: 109, SEQ ID NO: 111, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 117, SEQ ID NO: 119, SEQ ID NO: 122, SEQ ID NO: 126, SEQ ID NO: 132, SEQ ID NO: 134, SEQ ID NO: 135, SEQ ID NO: 136, SEQ ID NO: 138, SEQ ID NO: 139, SEQ ID NO: 140, and SEQ ID NO: 141.

Particularly preferred is a purified or isolated *H. pylori* cellular polypeptide or a fragment thereof, wherein the polypeptide is selected from the group consisting of SEQ ID NO: 85, SEQ ID NO: 88, SEQ ID NO: 93, SEQ ID NO: 96, SEQ ID NO: 97, SEQ

ID NO: 99, SEQ ID NO: 100, SEQ ID NO: 120, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 137, SEQ ID NO: 142, SEQ ID NO: 143, and SEQ ID NO: 146.

In another aspect, the invention pertains to any individual *H. pylori* polypeptide member or nucleic acid encoding such a member from the above-identified groups of *H. pylori* polypeptides.

In another aspect, the invention features nucleic acids capable of binding mRNA of *H. pylori*. Such nucleic acid is capable of acting as antisense nucleic acid to control the translation of mRNA of *H. pylori*. A further aspect features a nucleic acid which is capable of binding specifically to an *H. pylori* nucleic acid. These nucleic acids are also referred to herein as complements and have utility as probes and as capture reagents.

In another aspect, the invention features an expression system comprising an open reading frame corresponding to *H. pylori* nucleic acid. The nucleic acid further comprises a control sequence compatible with an intended host. The expression system is useful for making polypeptides corresponding to *H. pylori* nucleic acid.

In another aspect, the invention features a cell transformed with the expression system to produce *H. pylori* polypeptides.

In another aspect, the invention features a method of generating antibodies against *H. pylori* polypeptides which are capable of binding specifically to *H. pylori* polypeptides. Such antibodies have utility as reagents for immunoassays to evaluate the abundance and distribution of *H. pylori*-specific antigens.

In another aspect, the invention features a method of generating vaccines for immunizing an individual against *H. pylori*. The vaccination method includes: immunizing a subject with at least one *H. pylori* polypeptide according to the present invention, e.g., a surface or secreted polypeptide, or active portion thereof, and a pharmaceutically acceptable carrier. Such vaccines have therapeutic and/or prophylactic utilities.

In another aspect, the invention provides a method for generating a vaccine comprising a modified immunogenic *H. pylori* polypeptide, e.g., a surface or secreted polypeptide, or active portion thereof, and a pharmacologically acceptable carrier.

In another aspect, the invention features a method of evaluating a compound, e.g. a polypeptide, e.g., a fragment of a host cell polypeptide, for the ability to bind an *H. pylori* polypeptide. The method includes: contacting the candidate compound with an *H. pylori* polypeptide and determining if the compound binds or otherwise interacts with an *H. pylori* polypeptide. Compounds which bind *H. pylori* are candidates as activators or inhibitors of the bacterial life cycle. These assays can be performed *in vitro* or *in vivo*.

In another aspect, the invention features a method of evaluating a compound, e.g. a polypeptide, e.g., a fragment of a host cell polypeptide, for the ability to bind an *H. pylori* nucleic acid, e.g., DNA or RNA. The method includes: contacting the candidate compound with an *H. pylori* nucleic acid and determining if the compound binds or otherwise interacts with an *H. pylori* polypeptide. Compounds which bind *H. pylori* are candidates as activators or inhibitors of the bacterial life cycle. These assays can be performed *in vitro* or *in vivo*.

The invention features *H. pylori* polypeptides, preferably a substantially pure preparation of an *H. pylori* polypeptide, or a recombinant *H. pylori* polypeptide. In preferred embodiments: the polypeptide has biological activity; the polypeptide has an amino acid sequence at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% identical or homologous to an amino acid sequence of the invention contained in the Sequence Listing, preferably it has about 65% sequence identity with an amino acid sequence of the invention contained in the Sequence Listing, and most preferably it has about 92% to about 99% sequence identity with an amino acid sequence of the invention contained in the Sequence Listing; the polypeptide has an amino acid sequence essentially the same as an amino acid sequence of the invention contained in the Sequence Listing; the polypeptide is at least 5, 10, 20, 50, 100, or 150 amino acid residues in length; the polypeptide includes at least 5, preferably at least 10, more preferably at least 20, more preferably at least 50, 100, or 150 contiguous amino acid residues of the invention contained in the Sequence Listing. In yet another preferred embodiment, the amino acid sequence which differs in sequence identity by about 7% to about 8% from the *H. pylori* amino acid sequences of the invention contained in the Sequence Listing is also encompassed by the invention.

In preferred embodiments: the *H. pylori* polypeptide is encoded by a nucleic acid of the invention contained in the Sequence Listing, or by a nucleic acid having at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% homology with a nucleic acid of the invention contained in the Sequence Listing.

In a preferred embodiment, the subject *H. pylori* polypeptide differs in amino acid sequence at 1, 2, 3, 5, 10 or more residues from a sequence of the invention contained in the Sequence Listing. The differences, however, are such that the *H. pylori* polypeptide exhibits an *H. pylori* biological activity, e.g., the *H. pylori* polypeptide retains a biological activity of a naturally occurring *H. pylori* polypeptide.

In preferred embodiments, the polypeptide includes all or a fragment of an amino acid sequence of the invention contained in the Sequence Listing; fused, in reading frame, to additional amino acid residues, preferably to residues encoded by genomic

DNA 5' or 3' to the genomic DNA which encodes a sequence of the invention contained in the Sequence Listing.

In yet other preferred embodiments, the *H. pylori* polypeptide is a recombinant fusion protein having a first *H. pylori* polypeptide portion and a second polypeptide portion, e.g., a second polypeptide portion having an amino acid sequence unrelated to *H. pylori*. The second polypeptide portion can be, e.g., any of glutathione-S-transferase, a DNA binding domain, or a polymerase activating domain. In preferred embodiment the fusion protein can be used in a two-hybrid assay.

Polypeptides of the invention include those which arise as a result of alternative transcription events, alternative RNA splicing events, and alternative translational and postranslational events.

The invention also encompasses an immunogenic component which includes at least one *H. pylori* polypeptide in an immunogenic preparation; the immunogenic component being capable of eliciting an immune response specific for the *H. pylori* polypeptide, e.g., a humoral response, an antibody response, or a cellular response. In preferred embodiments, the immunogenic component comprises at least one antigenic determinant from a polypeptide of the invention contained in the Sequence Listing.

In another aspect, the invention provides a substantially pure nucleic acid having a nucleotide sequence which encodes an *H. pylori* polypeptide. In preferred embodiments: the encoded polypeptide has biological activity; the encoded polypeptide has an amino acid sequence at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% homologous to an amino acid sequence of the invention contained in the Sequence Listing; the encoded polypeptide has an amino acid sequence essentially the same as an amino acid sequence of the invention contained in the Sequence Listing; the encoded polypeptide is at least 5, 10, 20, 50, 100, or 150 amino acids in length; the encoded polypeptide comprises at least 5, preferably at least 10, more preferably at least 20, more preferably at least 50, 100, or 150 contiguous amino acids of the invention contained in the Sequence Listing.

In preferred embodiments: the nucleic acid of the invention is that contained in the Sequence Listing; the nucleic acid is at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% homologous with a nucleic acid sequence of the invention contained in the Sequence Listing.

In a preferred embodiment, the encoded *H. pylori* polypeptide differs (e.g., by amino acid substitution, addition or deletion of at least one amino acid residue) in amino acid sequence at 1, 2, 3, 5, 10 or more residues, from a sequence of the invention contained in the Sequence Listing. The differences, however, are such that: the *H.*

*pylori* encoded polypeptide exhibits a *H. pylori* biological activity, e.g., the encoded *H. pylori* enzyme retains a biological activity of a naturally occurring *H. pylori*.

In preferred embodiments, the encoded polypeptide includes all or a fragment of an amino acid sequence of the invention contained in the Sequence Listing; fused, in reading frame, to additional amino acid residues, preferably to residues encoded by genomic DNA 5' or 3' to the genomic DNA which encodes a sequence of the invention contained in the Sequence Listing.

In preferred embodiments, the subject *H. pylori* nucleic acid will include a transcriptional regulatory sequence, e.g. at least one of a transcriptional promoter or transcriptional enhancer sequence, operably linked to the *H. pylori* gene sequence, e.g., to render the *H. pylori* gene sequence suitable for expression in a recombinant host cell.

In yet a further preferred embodiment, the nucleic acid which encodes an *H. pylori* polypeptide of the invention, hybridizes under stringent conditions to a nucleic acid probe corresponding to at least 8 consecutive nucleotides of the invention contained in the Sequence Listing; more preferably to at least 12 consecutive nucleotides of the invention contained in the Sequence Listing; more preferably to at least 20 consecutive nucleotides of the invention contained in the Sequence Listing; more preferably to at least 40 consecutive nucleotides of the invention contained in the Sequence Listing.

In a preferred embodiment, the nucleic acid encodes a peptide which differs by at least one amino acid residue from the sequences of the invention contained in the Sequence Listing.

In a preferred embodiment, the nucleic acid differs by at least one nucleotide from a nucleotide sequence of the invention contained in the Sequence Listing which encodes amino acids of the invention contained in the Sequence Listing.

In another aspect, the invention encompasses: a vector including a nucleic acid which encodes an *H. pylori* polypeptide or an *H. pylori* polypeptide variant as described herein; a host cell transfected with the vector; and a method of producing a recombinant *H. pylori* polypeptide or *H. pylori* polypeptide variant; including culturing the cell, e.g., in a cell culture medium, and isolating the *H. pylori* or *H. pylori* polypeptide variant, e.g., from the cell or from the cell culture medium.

In another aspect, the invention features, a purified recombinant nucleic acid having at least 50%, 60%, 70%, 80%, 90%, 95%, 98%, or 99% homology with a sequence of the invention contained in the Sequence Listing.

The invention also provides a probe or primer which includes a substantially purified oligonucleotide. The oligonucleotide includes a region of nucleotide sequence which hybridizes under stringent conditions to at least 8 consecutive nucleotides of sense or antisense sequence of the invention contained in the Sequence Listing, or

naturally occurring mutants thereof. In preferred embodiments, the probe or primer further includes a label group attached thereto. The label group can be, e.g., a radioisotope, a fluorescent compound, an enzyme, and/or an enzyme co-factor. Preferably the oligonucleotide is at least 8 and less than 10, 20, 30, 50, 100, or 150 nucleotides in length.

The invention also provides an isolated *H. pylori* polypeptide which is encoded by a nucleic acid which hybridizes under stringent hybridization conditions to a nucleic acid contained in the Sequence Listing.

The invention further provides nucleic acids, e.g., RNA or DNA, encoding a polypeptide of the invention. This includes double stranded nucleic acids as well as coding and antisense single strands.

The *H. pylori* strain, from which genomic sequences have been sequenced, has been deposited in the American Type Culture Collection (ATCC # 55679; deposited by Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154) as strain HP-J99.

Included in the invention are: allelic variations; natural mutants; induced mutants; proteins encoded by DNA that hybridizes under high or low stringency conditions to a nucleic acid which encodes a polypeptide of the invention contained in the Sequence Listing (for definitions of high and low stringency see Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989, 6.3.1 - 6.3.6 and 6.4.1-6.4.10, hereby incorporated by reference); and, polypeptides specifically bound by antisera to *H. pylori* polypeptides, especially by antisera to an active site or binding domain of *H. pylori* polypeptide. The invention also includes fragments, preferably biologically active fragments. These and other polypeptides are also referred to herein as *H. pylori* polypeptide analogs or variants.

Putative functions have been determined for several of the *H. pylori* polypeptides of the invention, as shown in Table 1.

Accordingly, uses of the claimed *H. pylori* polypeptides based on these identified functions, as well as other functions as described herein, are also within the scope of the invention.

In addition, the present invention encompasses *H. pylori* polypeptides characterized as shown in Table 1 below, including: *H. pylori* cell envelope proteins, *H. pylori* secreted proteins, and *H. pylori* cellular proteins. Members of these groups were identified by BLAST homology searches and by searches for secretion signal or transmembrane protein motifs. Polypeptides related by significant homology to the polypeptides of Table 1 are also considered to be classified in the manner of the homologs shown in Table 1.

TABLE 1

| ORF_Name and Group | nt SeqID | aa SeqID |
|---|---|---|
| A. CELL ENVELOPE | | |
| A.1 Inner membrane proteins | | |
| 02ge11622_23494043_f1_6 | 3 | 76 |
| hp5p15212_13095752_c3_36 | 25 | 98 |
| 06ep30223_20173437_f1_37 | 48 | 121 |
| A.2 Outer membrane proteins | | |
| 05ee10816_14495437_f2_13 | 10 | 83 |
| A.2.1 Terminal phe residue | | |
| 06ep11509_35954752_f2_1 | 16 | 89 |
| 06ep10615_14495437_f3_47 | 45 | 118 |
| 03ae10804_14495437_c2_38 | 35 | 108 |
| 05ae30220_917200_c3_172 | 37 | 110 |
| 04cp11202_23646885_f2_26 | 7 | 80 |
| 05ep10815_16131925_c2_97 | 39 | 112 |
| 09cp61003_5860877_f2_23 | 55 | 128 |
| 09ae10512_48768_c3_67 | 18 | 91 |
| 09cp11003_5860877_f3_7 | 19 | 92 |
| hp6e12267_30478562_f3_33 | 28 | 101 |
| 06cp30603_34174212_c3_71 | 30 | 103 |
| 09cp10224_1962590_f3_31 | 52 | 125 |
| 09cp61003_30478562_c3_106 | 54 | 127 |
| 11ae80818_10553192_f2_16 | 56 | 129 |
| 11ee11408_10584582_c3_51 | 58 | 131 |
| A.2.2 Terminal phe residue and C-terminal tyrosine cluster | | |
| 01ae12001_116018_c2_40 | 1 | 74 |
| 06ap10609_116018_c3_50 | 42 | 115 |
| 06cp30603_4687507_f1_9 | 14 | 87 |
| 06cp30603_4687507_f1_7 | 43 | 116 |
| 05ee10816_36126938_f3_16 | 11 | 84 |
| 01cp20708_4960952_c1_43 | 71 | 144 |
| A.3 Via homolgy | | |
| 07ap80601_5083193_f3_8 | 17 | 90 |
| 11ap20714_4797137_f3_45 | 57 | 130 |
| A.4 Other cell envelope proteins | | |
| 04ap12016_25501501_f1_1 | 5 | 78 |
| 04cp11202_20415937_f2_25 | 6 | 79 |
| 04ee11108_3906963_f1_7 | 8 | 81 |
| 29ep10720_25501501_c2_33 | 21 | 94 |
| B. SECRETED PROTEINS | | |
| hp3e10342_22448587_c2_15 | 72 | 145 |
| hp5p15212_24276587_f1_2 | 32 | 105 |
| 09ce10413_35336707_f2_9 | 51 | 124 |

| | | |
|---|---|---|
| 01ae12001_32462543_c2_43 | 2 | 75 |
| 03ee11215_1416312_c3_35 | 4 | 77 |
| 05ae30220_14570443_c2_94 | 9 | 82 |
| 06cp30603_2772578_c1_46 | 13 | 86 |
| 29ep10720_289077_f2_12 | 22 | 95 |
| 03ee11215_22542803_f1_7 | 29 | 102 |
| 09ae10512_3166040_c1_40 | 31 | 104 |
| 01ce11104_10742963_c2_12 | 33 | 106 |
| 02ge10116_36335436_f3_66 | 34 | 107 |
| 04ep41903_11876461_f1_4 | 36 | 109 |
| 05ce10208_23631292_f1_6 | 38 | 111 |
| 05ep10815_22447252_c3_110 | 40 | 113 |
| 05ep10815_30283516_c3_109 | 41 | 114 |
| 06ee30709_33851038_c3_30 | 44 | 117 |
| 06ep11202_21687842_c3_35 | 46 | 119 |
| 06ep30223_2774062_f1_33 | 49 | 122 |
| 09cp10713_23912707_c1_26 | 53 | 126 |
| 11ee11408_4882318_f3_24 | 59 | 132 |
| hp4e13394_5908553_f1_1 | 61 | 134 |
| hp4e53394_1416312_c3_119 | 62 | 135 |
| hp5e15211_24328910_c3_38 | 63 | 136 |
| hp6p10606_23493756_c1_21 | 65 | 138 |
| hp6p22217_23564012_f1_5 | 66 | 139 |
| hp6p22217_272058_f1_2 | 67 | 140 |
| hp6p22217_2922143_f2_9 | 68 | 141 |
| C. OTHER CELLULAR PROTEINS | | |
| 06ap11119_14726542_f3_21 | 12 | 85 |
| 06ee10709_6136430_c1_11 | 15 | 88 |
| 12ap10605_14094816_c1_5 | 20 | 93 |
| hp2p10272_34042518_f1_2 | 23 | 96 |
| hp5e15211_25411557_c1_22 | 24 | 97 |
| hp5p15641_3907968_f1_3 | 26 | 99 |
| hp6e10967_657638_f3_9 | 27 | 100 |
| 06ep11202_4569693_c2_28 | 47 | 120 |
| 06ep30223_3930468_c1_110 | 50 | 123 |
| hp2e10911_960952_c2_86 | 60 | 133 |
| hp6p10509_14642217_c2_17 | 64 | 137 |
| hp6p80503_20964382_f2_11 | 69 | 142 |
| hp7e10192_5917593_f1_2 | 70 | 143 |
| hp6p10509_14642217_c3_25 | 73 | 146 |

[In Table 1, "nt" represents nucleotide Seq. ID number and "aa" represents amino acid Seq. ID number]

Definitions

The terms "purified polypeptide" and "isolated polypeptide" and "a substantially pure preparation of a polypeptide" are used interchangeably herein and, as used herein,

from other proteins, lipids, and nucleic acids with which it naturally occurs. Preferably, the polypeptide is also separated from substances, e.g., antibodies or gel matrix, e.g., polyacrylamide, which are used to purify it. Preferably, the polypeptide constitutes at least 10, 20, 50 70, 80 or 95% dry weight of the purified preparation. Preferably, the preparation contains: sufficient polypeptide to allow protein sequencing; at least 1, 10, or 100 μg of the polypeptide; at least 1, 10, or 100 mg of the polypeptide. Furthermore, the terms "purified polypeptide" and "isolated polypeptide" and "a substantially pure preparation of a polypeptide," as used herein, refer to both a polypeptide obtained from nature or produced by recombinant DNA techniques as described herein.

For example, an "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the *H. pylori* protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of *H. pylori* protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of *H. pylori* protein having less than about 30% (by dry weight) of non-*H. pylori* protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-*H. pylori* protein, still more preferably less than about 10% of non-*H. pylori* protein, and most preferably less than about 5% non-*H. pylori* protein. When the *H. pylori* protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of *H. pylori* protein in which the protein is separated from chemical precusors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of *H. pylori* protein having less than about 30% (by dry weight) of chemical precursors or non-*H. pylori* chemicals, more preferably less than about 20% chemical precursors or non-*H. pylori* chemicals, still more preferably less than about 10% chemical precursors or non-*H. pylori* chemicals, and most preferably less than about 5% chemical precursors or non-*H. pylori* chemicals.

A purified preparation of cells refers to, in the case of plant or animal cells, an *in vitro* preparation of cells and not an entire intact plant or animal. In the case of cultured

cells or microbial cells, it consists of a preparation of at least 10% and more preferably 50% of the subject cells.

A purified or isolated or a substantially pure nucleic acid, e.g., a substantially pure DNA, (are terms used interchangeably herein) is a nucleic acid which is one or both of the following: not immediately contiguous with both of the coding sequences with which it is immediately contiguous (i.e., one at the 5' end and one at the 3' end) in the naturally-occurring genome of the organism from which the nucleic acid is derived; or which is substantially free of a nucleic acid with which it occurs in the organism from which the nucleic acid is derived. The term includes, for example, a recombinant DNA which is incorporated into a vector, e.g., into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other DNA sequences. Substantially pure DNA also includes a recombinant DNA which is part of a hybrid gene encoding additional *H. pylori* DNA sequence.

A "contig" as used herein is a nucleic acid representing a continuous stretch of genomic sequence of an organism.

An "open reading frame", also referred to herein as ORF, is a region of nucleic acid which encodes a polypeptide. This region may represent a portion of a coding sequence or a total sequence and can be determined from a stop to stop codon or from a start to stop codon.

As used herein, a "coding sequence" is a nucleic acid which is transcribed into messenger RNA and/or translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the five prime terminus and a translation stop code at the three prime terminus. A coding sequence can include but is not limited to messenger RNA, synthetic DNA, and recombinant nucleic acid sequences.

A "complement" of a nucleic acid as used herein referes to an anti-parallel or antisense sequence that participates in Watson-Crick base-pairing with the original sequence.

A "gene product" is a protein or structural RNA which is specifically encoded by a gene.

As used herein, the term "probe" refers to a nucleic acid, peptide or other chemical entity which specifically binds to a molecule of interest. Probes are often associated with or capable of associating with a label. A label is a chemical moiety capable of detection. Typical labels comprise dyes, radioisotopes, luminescent and chemiluminescent moieties, fluorophores, enzymes, precipitating agents, amplification

sequences, and the like. Similarly, a nucleic acid, peptide or other chemical entity which specifically binds to a molecule of interest and immobilizes such molecule is referred herein as a "capture ligand". Capture ligands are typically associated with or capable of associating with a support such as nitro-cellulose, glass, nylon membranes, beads, particles and the like. The specificity of hybridization is dependent on conditions such as the base pair composition of the nucleotides, and the temperature and salt concentration of the reaction. These conditions are readily discernable to one of ordinary skill in the art using routine experimentation.

Homologous refers to the sequence similarity or sequence identity between two polypeptides or between two nucleic acid molecules. When a position in both of the two compared sequences is occupied by the same base or amino acid monomer subunit, e.g., if a position in each of two DNA molecules is occupied by adenine, then the molecules are homologous at that position. The percent of homology between two sequences is a function of the number of matching or homologous positions shared by the two sequences divided by the number of positions compared x 100. For example, if 6 of 10 of the positions in two sequences are matched or homologous then the two sequences are 60% homologous. By way of example, the DNA sequences ATTGCC and TATGGC share 50% homology. Generally, a comparison is made when two sequences are aligned to give maximum homology.

Nucleic acids are hybridizable to each other when at least one strand of a nucleic acid can anneal to the other nucleic acid under defined stringency conditions. Stringency of hybridization is determined by: (a) the temperature at which hybridization and/or washing is performed; and (b) the ionic strength and polarity of the hybridization and washing solutions. Hybridization requires that the two nucleic acids contain complementary sequences; depending on the stringency of hybridization, however, mismatches may be tolerated. Typically, hybridization of two sequences at high stingency (such as, for example, in a solution of 0.5X SSC, at 65° C) requires that the sequences be essentially completely homologous. Conditions of intermediate stringency (such as, for example, 2X SSC at 65 ° C) and low stringency (such as, for example 2X SSC at 55° C), require correspondingly less overall complementarity between the hybridizing sequences. (1X SSC is 0.15 M NaCl, 0.015 M Na citrate). A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C.

The terms peptides, proteins, and polypeptides are used interchangeably herein.

As used herein, the term "surface protein" refers to all surface accessible proteins, e.g. inner and outer membrane proteins, proteins adhering to the cell wall, and secreted proteins.

A polypeptide has *H. pylori* biological activity if it has one, two and preferably more of the following properties: (1) if when expressed in the course of an *H. pylori* infection, it can promote, or mediate the attachment of *H. pylori* to a cell; (2) it has an enzymatic activity, structural or regulatory function characteristic of an *H. pylori* protein; (3) the gene which encodes it can rescue a lethal mutation in an *H. pylori* gene; (4) or it is immunogenic in a subject. A polypeptide has biological activity if it is an antagonist, agonist, or super-agonist of a polypeptide having one of the above-listed properties.

A biologically active fragment or analog is one having an *in vivo* or *in vitro* activity which is characteristic of the *H. pylori* polypeptides of the invention contained in the Sequence Listing, or of other naturally occurring *H. pylori* polypeptides, e.g., one or more of the biological activities described herein. Especially preferred are fragments which exist *in vivo*, e.g., fragments which arise from post transcriptional processing or which arise from translation of alternatively spliced RNA's. Fragments include those expressed in native or endogenous cells as well as those made in expression systems, e.g., in CHO cells. Because peptides such as *H. pylori* polypeptides often exhibit a range of physiological properties and because such properties may be attributable to different portions of the molecule, a useful *H. pylori* fragment or *H. pylori* analog is one which exhibits a biological activity in any biological assay for *H. pylori* activity. Most preferably the fragment or analog possesses 10%, preferably 40%, more preferably 60%, 70%, 80% or 90% or greater of the activity of *H. pylori*, in any *in vivo* or *in vitro* assay.

Analogs can differ from naturally occurring *H. pylori* polypeptides in amino acid sequence or in ways that do not involve sequence, or both. Non-sequence modifications include changes in acetylation, methylation, phosphorylation, carboxylation, or glycosylation. Preferred analogs include *H. pylori* polypeptides (or biologically active fragments thereof) whose sequences differ from the wild-type sequence by one or more conservative amino acid substitutions or by one or more non-conservative amino acid substitutions, deletions, or insertions which do not substantially diminish the biological activity of the *H. pylori* polypeptide. Conservative substitutions typically include the substitution of one amino acid for another with similar characteristics, e.g., substitutions within the following groups: valine, glycine; glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. Other conservative substitutions can be made in view of the table below.

**TABLE 2**

CONSERVATIVE AMINO ACID REPLACEMENTS

| For Amino Acid | Code | Replace with any of |
|---|---|---|
| Alanine | A | D-Ala, Gly, beta-Ala, L-Cys, D-Cys |
| Arginine | R | D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn |
| Asparagine | N | D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln |
| Aspartic Acid | D | D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln |
| Cysteine | C | D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr |
| Glutamine | Q | D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp |
| Glutamic Acid | E | D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln |
| Glycine | G | Ala, D-Ala, Pro, D-Pro, β-Ala, Acp |
| Isoleucine | I | D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met |
| Leucine | L | D-Leu, Val, D-Val, Leu, D-Leu, Met, D-Met |
| Lysine | K | D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn |
| Methionine | M | D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val |
| Phenylalanine | F | D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-phenylproline, cis-3,4, or 5-phenylproline |
| Proline | P | D-Pro, L-I-thioazolidine-4-carboxylic acid, D-or L-1-oxazolidine-4-carboxylic acid |
| Serine | S | D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys |
| Threonine | T | D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val |
| Tyrosine | Y | D-Tyr, Phe, D-Phe, L-Dopa, His, D-His |
| Valine | V | D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met |

Other analogs within the invention are those with modifications which increase peptide stability; such analogs may contain, for example, one or more non-peptide bonds (which replace the peptide bonds) in the peptide sequence. Also included are: analogs that include residues other than naturally occurring L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids, e.g., β or γ amino acids; and cyclic analogs.

As used herein, the term "fragment", as applied to an *H. pylori* analog, will ordinarily be at least about 20 residues, more typically at least about 40 residues, preferably at least about 60 residues in length. Fragments of *H. pylori* polypeptides can be generated by methods known to those skilled in the art. The ability of a candidate fragment to exhibit a biological activity of *H. pylori* polypeptide can be assessed by methods known to those skilled in the art as described herein. Also included are *H. pylori* polypeptides containing residues that are not required for biological activity of the peptide or that result from alternative mRNA splicing or alternative protein processing events.

An "immunogenic component" as used herein is a moiety, such as an *H. pylori* polypeptide, analog or fragment thereof, that is capable of eliciting a humoral and/or cellular immune response in a host animal alone or in combination with an adjuvant.

An "antigenic component" as used herein is a moiety, such as an *H. pylori* polypeptide, analog or fragment thereof, that is capable of binding to a specific antibody with sufficiently high affinity to form a detectable antigen-antibody complex.

As used herein, the term "transgene" means a nucleic acid (encoding, e.g., one or more polypeptides), which is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the cell's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal expression of the selected nucleic acid, all operably linked to the selected nucleic acid, and may include an enhancer sequence.

As used herein, the term "transgenic cell" refers to a cell containing a transgene.

As used herein, a "transgenic animal" is any animal in which one or more, and preferably essentially all, of the cells of the animal includes a transgene. The transgene can be introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by a process of transformation of competent cells or by microinjection or by infection with a

recombinant virus. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA.

The term "antibody" as used herein is intended to include fragments thereof which are specifically reactive with *H. pylori* polypeptides.

As used herein, the term "cell-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in specific cells of a tissue. The term also covers so-called "leaky" promoters, which regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well.

Misexpression, as used herein, refers to a non-wild type pattern of gene expression. It includes: expression at non-wild type levels, i.e., over or under expression; a pattern of expression that differs from wild type in terms of the time or stage at which the gene is expressed, e.g., increased or decreased expression (as compared with wild type) at a predetermined developmental period or stage; a pattern of expression that differs from wild type in terms of decreased expression (as compared with wild type) in a predetermined cell type or tissue type; a pattern of expression that differs from wild type in terms of the splicing size, amino acid sequence, post-transitional modification, or biological activity of the expressed polypeptide; a pattern of expression that differs from wild type in terms of the effect of an environmental stimulus or extracellular stimulus on expression of the gene, e.g., a pattern of increased or decreased expression (as compared with wild type) in the presence of an increase or decrease in the strength of the stimulus.

As used herein, "host cells" and other such terms denoting microorganisms or higher eukaryotic cell lines cultured as unicellular entities refers to cells which can become or have been used as recipients for a recombinant vector or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood by individuals skilled in the art that the progeny of a single parental cell may not necessarily be completely identical in genomic or total DNA compliment to the original parent, due to accident or deliberate mutation.

As used herein, the term "control sequence" refers to a nucleic acid having a base sequence which is recognized by the host organism to effect the expression of encoded sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include a promoter, ribosomal binding site, terminators, and in some cases operators; in eukaryotes, generally such control sequences include promoters, terminators and in some instances, enhancers. The term control sequence is intended to include at a

minimum, all components whose presence is necessary for expression, and may also include additional components whose presence is advantageous, for example, leader sequences.

As used herein, the term "operably linked" refers to sequences joined or ligated to function in their intended manner. For example, a control sequence is operably linked to coding sequence by ligation in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequence and host cell.

The metabolism of a substance, as used herein, means any aspect of the, expression, function, action, or regulation of the substance. The metabolism of a substance includes modifications, e.g., covalent or non-covalent modifications of the substance. The metabolism of a substance includes modifications, e.g., covalent or non-covalent modification, the substance induces in other substances. The metabolism of a substance also includes changes in the distribution of the substance. The metabolism of a substance includes changes the substance induces in the distribution of other substances.

A "sample" as used herein refers to a biological sample, such as, for example, tissue or fluid isloated from an individual (including without limitation plasma, serum, cerebrospinal fluid, lymph, tears, saliva and tissue sections) or from *in vitro* cell culture constituents, as well as samples from the environment.

The practice of the invention will employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. *See e.g.*, Sambrook, Fritsch, and Maniatis, *Molecular Cloning; Laboratory Manual* 2nd ed. (1989); *DNA Cloning*, Volumes I and II (D.N Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed, 1984); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds. 1984); the series, *Methods in Enzymology* (Academic Press, Inc.), particularly Vol. 154 and Vol. 155 (Wu and Grossman, eds.) and *PCR-A Practical Approach* (McPherson, Quirke, and Taylor, eds., 1991).

I. Isolation of Nucleic Acids of *H. pylori* and Uses Therefor

*H. pylori* Genomic Sequence

This invention provides nucleotide sequences of the genome of *H. pylori* which thus comprises a DNA sequence library of *H. pylori* genomic DNA. The detailed description that follows provides nucleotide sequences of *H. pylori*, and also describes how the sequences were obtained and how ORFs and protein-coding sequences were

identified. Also described are methods of using the disclosed *H. pylori* sequences in methods including diagnostic and therapeutic applications. Furthermore, the library can be used as a database for identification and comparison of medically important sequences in this and other strains of *H. pylori*.

To determine the genomic sequence of *H. pylori*, DNA was isolated from a strain of *H. pylori* (ATCC # 55679; deposited by Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154) and mechanically sheared by nebulization to a median size of 2 kb. Following size fractionation by gel electrophoresis, the fragments were blunt-ended, ligated to adapter oligonucleotides, and cloned into each of 20 different pMPX vectors (Rice et al., abstracts of Meeting of Genome Mapping and Sequencing, Cold Spring Harbor, NY, 5/11-5/15, 1994, p. 225) to construct a series of "shotgun" subclone libraries.

DNA sequencing was achieved using multiplex sequencing procedures essentially as disclosed in Church et al., 1988, *Science* 240:185; U.S. Patents No. 4,942,124 and 5,149,625). DNA was extracted from pooled cultures and subjected to chemical or enzymatic sequencing. Sequencing reactions were resolved by electrophoresis, and the products were transferred and covalently bound to nylon membranes. Finally, the membranes were sequentially hybridized with a series of labelled oligonucleotides complimentary to "tag" sequences present in the different shotgun cloning vectors. In this manner, a large number of sequences could be obtained from a single set of sequencing reactions. The cloning and sequencing procedures are described in more detail in the Exemplification.

Individual sequence reads obtained in this manner were assembled using the FALCON™ program (Church *et al.*, 1994, *Automated DNA Sequencing and Analysis*, J.C. Venter, ed., Academic Press) and PHRAP (P. Green, Abstracts of DOE Human Genome Program Contractor-Grantee Workshop V, Jan. 1996, p.157). The average contig length was about 3-4 kb.

A variety of approaches are used to order the contigs so as to obtain a continuous sequence representing the entire *H. pylori* genome. Synthetic oligonucleotides are designed that are complementary to sequences at the end of each contig. These oligonucleotides may be hybridized to libaries of *H. pylori* genomic DNA in, for example, lambda phage vectors or plasmid vectors to identify clones that contain sequences corresponding to the junctional regions between individual contigs. Such clones are then used to isolate template DNA and the same oligonucleotides are used as primers in polymerase chain reaction (PCR) to amplify junctional fragments, the nucleotide sequence of which is then determined.

The *H. pylori* sequences were analyzed for the presence of open reading frames (ORFs) comprising at least 180 nucleotides. As a result of the analysis of ORFs based on stop-to-stop codon reads, it should be understood that these ORFs may not correspond to the ORF of a naturally-occurring *H. pylori* polypeptide. These ORFs may contain start codons which indicate the initiation of protein synthesis of a naturally-occurring *H. pylori* polypeptide. Such start codons within the ORFs provided herein can be identified by those of ordinary skill in the relevant art, and the resulting ORF and the encoded *H. pylori* polypeptide is within the scope of this invention. For example, within the ORFs a codon such as AUG or GUG (encoding methionine or valine) which is part of the initiation signal for protein synthesis can be identified and the ORF modified to correspond to a naturally-occurring *H. pylori* polypeptide. The predicted coding regions were defined by evaluating the coding potential of such sequences with the program GENEMARK™ (Borodovsky and McIninch, 1993, *Comp. Chem.* 17:123).

Other *H. pylori* Nucleic Acids

The nucleic acids of this invention may be obtained directly from the DNA of the above referenced *H. pylori* strain by using the polymerase chain reaction (PCR). See "*PCR, A Practical Approach*" (McPherson, Quirke, and Taylor, eds., IRL Press, Oxford, UK, 1991) for details about the PCR. High fidelity PCR can be used to ensure a faithful DNA copy prior to expression. In addition, the authenticity of amplified products can be checked by conventional sequencing methods. Clones carrying the desired sequences described in this invention may also be obtained by screening the libraries by means of the PCR or by hybridization of synthetic oligonucleotide probes to filter lifts of the library colonies or plaques as known in the art (see, e.g., Sambrook et al., *Molecular Cloning, A Laboratory Manual* 2nd edition, 1989, Cold Spring Harbor Press, NY).

It is also possible to obtain nucleic acids encoding *H. pylori* polypeptides from a cDNA library in accordance with protocols herein described. A cDNA encoding an *H. pylori* polypeptide can be obtained by isolating total mRNA from an appropriate strain. Double stranded cDNAs can then be prepared from the total mRNA. Subsequently, the cDNAs can be inserted into a suitable plasmid or viral (e.g., bacteriophage) vector using any one of a number of known techniques. Genes encoding *H. pylori* polypeptides can also be cloned using established polymerase chain reaction techniques in accordance with the nucleotide sequence information provided by the invention. The nucleic acids of the invention can be DNA or RNA. Preferred nucleic acids of the invention are contained in the Sequence Listing.

The nucleic acids of the invention can also be chemically synthesized using standard techniques. Various methods of chemically synthesizing polydeoxynucleotides

are known, including solid-phase synthesis which, like peptide synthesis, has been fully automated in commercially available DNA synthesizers (See e.g., Itakura et al. U.S. Patent No. 4,598,049; Caruthers et al. U.S. Patent No. 4,458,066; and Itakura U.S. Patent Nos. 4,401,796 and 4,373,071, incorporated by reference herein).

Nucleic acids isolated or synthesized in accordance with features of the present invention are useful, by way of example, without limitation, as probes, primers, capture ligands, antisense genes and for developing expression systems for the synthesis of proteins and peptides corresponding to such sequences. As probes, primers, capture ligands and antisense agents, the nucleic acid normally consists of all or part (approximately twenty or more nucleotides for specificity as well as the ability to form stable hybridization products) of the nucleic acids of the invention contained in the Sequence Listing. These uses are described in further detail below.

Probes

A nucleic acid isolated or synthesized in accordance with the sequence of the invention contained in the Sequence Listing can be used as a probe to specifically detect *H. pylori*. With the sequence information set forth in the present application, sequences of twenty or more nucleotides are identified which provide the desired inclusivity and exclusivity with respect to *H. pylori*, and extraneous nucleic acids likely to be encountered during hybridization conditions. More preferably, the sequence will comprise at least twenty to thirty nucleotides to convey stability to the hybridization product formed between the probe and the intended target molecules.

Sequences larger than 1000 nucleotides in length are difficult to synthesize but can be generated by recombinant DNA techniques. Individuals skilled in the art will readily recognize that the nucleic acids, for use as probes, can be provided with a label to facilitate detection of a hybridization product.

Nucleic acid isolated and synthesized in accordance with the sequence of the invention contained in the Sequence Listing can also be useful as probes to detect homologous regions (especially homologous genes) of other *Helicobacter* species using appropriate stringency hybridization conditions as described herein.

Capture Ligand

For use as a capture ligand, the nucleic acid selected in the manner described above with respect to probes, can be readily associated with a support. The manner in which nucleic acid is associated with supports is well known. Nucleic acid having twenty or more nucleotides in a sequence of the invention contained in the Sequence Listing have utility to separate *H. pylori* nucleic acid from the nucleic acid of each other and other organisms. Nucleic acid having twenty or more nucleotides in a sequence of the invention contained in the Sequence Listing can also have utility to separate other

*Helicobacter* species from each other and from other organisms. Preferably, the sequence will comprise at least twenty nucleotides to convey stability to the hybridization product formed between the probe and the intended target molecules. Sequences larger than 1000 nucleotides in length are difficult to synthesize but can be generated by recombinant DNA techniques.

Primers

Nucleic acid isolated or synthesized in accordance with the sequences described herein have utility as primers for the amplification of *H. pylori* nucleic acid. These nucleic acids may also have utility as primers for the amplification of nucleic acids in other *Helicobacter* species. With respect to polymerase chain reaction (PCR) techniques, nucleic acid sequences of $\geq$ 10-15 nucleotides of the invention contained in the Sequence Listing have utility in conjunction with suitable enzymes and reagents to create copies of *H. pylori* nucleic acid. More preferably, the sequence will comprise twenty or more nucleotides to convey stability to the hybridization product formed between the primer and the intended target molecules. Binding conditions of primers greater than 100 nucleotides are more difficult to control to obtain specificity. High fidelity PCR can be used to ensure a faithful DNA copy prior to expression. In addition, amplified products can be checked by conventional sequencing methods.

The copies can be used in diagnostic assays to detect specific sequences, including genes from *H. pylori* and/or other *Helicobacter* species. The copies can also be incorporated into cloning and expression vectors to generate polypeptides corresponding to the nucleic acid synthesized by PCR, as is described in greater detail herein.

Antisense

Nucleic acid or nucleic acid-hybridizing derivatives isolated or synthesized in accordance with the sequences described herein have utility as antisense agents to prevent the expression of *H. pylori* genes. These sequences also have utility as antisense agents to prevent expression of genes of other *Helicobacter* species.

In one embodiment, nucleic acid or derivatives corresponding to *H. pylori* nucleic acids is loaded into a suitable carrier such as a liposome or bacteriophage for introduction into bacterial cells. For example, a nucleic acid having twenty or more nucleotides is capable of binding to bacteria nucleic acid or bacteria messenger RNA. Preferably, the antisense nucleic acid is comprised of 20 or more nucleotides to provide necessary stability of a hybridization product of non-naturally occurring nucleic acid and bacterial nucleic acid and/or bacterial messenger RNA. Nucleic acid having a sequence greater than 1000 nucleotides in length is difficult to synthesize but can be generated by recombinant DNA techniques. Methods for loading antisense nucleic acid in liposomes

is known in the art as exemplified by U.S. Patent 4,241,046 issued December 23, 1980 to Papahadjopoulos et al.

II. Expression of *H. pylori* Nucleic Acids

Nucleic acid isolated or synthesized in accordance with the sequences described herein have utility to generate polypeptides. The nucleic acid of the invention exemplified in the Sequence Listing or fragments of said nucleic acid encoding active portions of *H. pylori* polypeptides can be cloned into suitable vectors or used to isolate nucleic acid. The isolated nucleic acid is combined with suitable DNA linkers and cloned into a suitable vector.

The function of a specific gene or operon can be ascertained by expression in a bacterial strain under conditions where the activity of the gene product(s) specified by the gene or operon in question can be specifically measured. Alternatively, a gene product may be produced in large quantities in an expressing strain for use as an antigen, an industrial reagent, for structural studies, etc. This expression can be accomplished in a mutant strain which lacks the activity of the gene to be tested, or in a strain that does not produce the same gene product(s). This includes, but is not limited to other *Helicobacter* strains, or other bacterial strains such as *E. coli*, *Norcardia*, *Corynebacterium*, *Campylobacter*, and *Streptomyces* species. In some cases the expression host will utilize the natural *Helicobacter* promoter whereas in others, it will be necessary to drive the gene with a promoter sequence derived from the expressing organism (e.g., an *E. coli* beta-galactosidase promoter for expression in *E. coli*).

To express a gene product using the natural *H. pylori* promoter, a procedure such as the following can be used. A restriction fragment containing the gene of interest, together with its associated natural promoter element and regulatory sequences (identified using the DNA sequence data) is cloned into an appropriate recombinant plasmid containing an origin of replication that functions in the host organism and an appropriate selectable marker. This can be accomplished by a number of procedures known to those skilled in the art. It is most preferably done by cutting the plasmid and the fragment to be cloned with the same restriction enzyme to produce compatible ends that can be ligated to join the two pieces together. The recombinant plasmid is introduced into the host organism by, for example, electroporation and cells containing the recombinant plasmid are identified by selection for the marker on the plasmid. Expression of the desired gene product is detected using an assay specific for that gene product.

In the case of a gene that requires a different promoter, the body of the gene (coding sequence) is specifically excised and cloned into an appropriate expression

plasmid. This subcloning can be done by several methods, but is most easily accomplished by PCR amplification of a specific fragment and ligation into an expression plasmid after treating the PCR product with a restriction enzyme or exonuclease to create suitable ends for cloning.

A suitable host cell for expression of a gene can be any procaryotic or eucaryotic cell. For example, an *H. pylori* polypeptide can be expressed in bacterial cells such as *E. coli*, insect cells (baculovirus), yeast, or mammalian cells such as Chinese hamster ovary cell (CHO). Other suitable host cells are known to those skilled in the art.

Expression in eucaryotic cells such as mammalian, yeast, or insect cells can lead to partial or complete glycosylation and/or formation of relevant inter- or intra-chain disulfide bonds of a recombinant peptide product. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari. et al., (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al., (1987) *Gene* 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA). Baculovirus vectors available for expression of proteins in cultured insect cells (SF 9 cells) include the pAc series (Smith et al., (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow, V.A., and Summers, M.D., (1989) *Virology* 170:31-39). Generally, COS cells (Gluzman, Y., (1981) Cell 23:175-182) are used in conjunction with such vectors as pCDM 8 (Aruffo, A. and Seed, B., (1987) *Proc. Natl. Acad. Sci. USA* 84:8573-8577) for transient amplification/expression in mammalian cells, while CHO (dhfr$^-$ Chinese Hamster Ovary) cells are used with vectors such as pMT2PC (Kaufman et al. (1987), *EMBO J.* 6:187-195) for stable amplification/expression in mammalian cells. Vector DNA can be introduced into mammalian cells via conventional techniques such as calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, or electroporation. Suitable methods for transforming host cells can be found in Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory press (1989)), and other laboratory textbooks.

Expression in procaryotes is most often carried out in *E. coli* with either fusion or non-fusion inducible expression vectors. Fusion vectors usually add a number of $NH_2$ terminal amino acids to the expressed target gene. These $NH_2$ terminal amino acids often are referred to as a reporter group. Such reporter groups usually serve two purposes: 1) to increase the solubility of the target recombinant protein; and 2) to aid in the purification of the target recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the reporter group and the target recombinant protein to enable separation of the target recombinant protein from the reporter group subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition

sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Amrad Corp., Melbourne, Australia), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase, maltose E binding protein, or protein A, respectively, to the target recombinant protein. A preferred reporter group is poly(His), which may be fused to the amino or carboxy terminus of the protein and which renders the recombinant fusion protein easily purifiable by metal chelate chromatography.

Inducible non-fusion expression vectors include pTrc (Amann et al., (1988) Gene 69:301-315) and pET11d (Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 60-89). While target gene expression relies on host RNA polymerase transcription from the hybrid trp-lac fusion promoter in pTrc, expression of target genes inserted into pET11d relies on transcription from the T7 gn10-lac 0 fusion promoter mediated by coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 under the transcriptional control of the lacUV 5 promoter.

For example, a host cell transfected with a nucleic acid vector directing expression of a nucleotide sequence encoding an *H. pylori* polypeptide can be cultured under appropriate conditions to allow expression of the polypeptide to occur. The polypeptide may be secreted and isolated from a mixture of cells and medium containing the peptide. Alternatively, the polypeptide may be retained cytoplasmically and the cells harvested, lysed and the protein isolated. A cell culture includes host cells, media and other byproducts. Suitable media for cell culture are well known in the art. Polypeptides of the invention can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for such polypeptides. Additionally, in many situations, polypeptides can be produced by chemical cleavage of a native protein (e.g., tryptic digestion) and the cleavage products can then be purified by standard techniques.

In the case of membrane bound proteins, these can be isolated from a host cell by contacting a membrane-associated protein fraction with a detergent forming a solubilized complex, where the membrane-associated protein is no longer entirely embedded in the membrane fraction and is solubilized at least to an extent which allows it to be chromatographically isolated from the membrane fraction. Several different criteria are used for choosing a detergent suitable for solubilizing these complexes. For example, one property considered is the ability of the detergent to solubilize the *H.*

*pylori* protein within the membrane fraction at minimal denaturation of the membrane-associated protein allowing for the activity or functionality of the membrane-associated protein to return upon reconstitution of the protein. Another property considered when selecting the detergent is the critical micelle concentration (CMC) of the detergent in that the detergent of choice preferably has a high CMC value allowing for ease of removal after reconstitution. A third property considered when selecting a detergent is the hydrophobicity of the detergent. Typically, membrane-associated proteins are very hydrophobic and therefore detergents which are also hydrophobic, e.g., the triton series, would be useful for solubilizing the hydrophobic proteins. Another property important to a detergent can be the capability of the detergent to remove the *H. pylori* protein with minimal protein-protein interaction facilitating further purification. A fifth property of the detergent which should be considered is the charge of the detergent. For example, if it is desired to use ion exchange resins in the purification process then preferably detergent should be an uncharged detergent. Chromatographic techniques which can be used in the final purification step are known in the art and include hydrophobic interaction, lectin affinity, ion exchange, dye affinity and immunoaffinity.

One strategy to maximize recombinant *H. pylori* peptide expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 119-128). Another strategy would be to alter the nucleic acid encoding an *H. pylori* peptide to be inserted into an expression vector so that the individual codons for each amino acid would be those preferentially utilized in highly expressed *E. coli* proteins (Wada et al., (1992) *Nuc. Acids Res.* 20:2111-2118). Such alteration of nucleic acids of the invention can be carried out by standard DNA synthesis techniques.

The nucleic acids of the invention can also be chemically synthesized using standard techniques. Various methods of chemically synthesizing polydeoxynucleotides are known, including solid-phase synthesis which, like peptide synthesis, has been fully automated in commercially available DNA synthesizers (See, e.g., Itakura et al. U.S. Patent No. 4,598,049; Caruthers et al. U.S. Patent No. 4,458,066; and Itakura U.S. Patent Nos. 4,401,796 and 4,373,071, incorporated by reference herein).

III. *H. pylori* Polypeptides

This invention encompasses isolated *H. pylori* polypeptides encoded by the disclosed *H. pylori* genomic sequences, including the polypeptides of the invention contained in the Sequence Listing. Polypeptides of the invention are preferably at least 5 amino acid residues in length. Using the DNA sequence information provided herein,

the amino acid sequences of the polypeptides encompassed by the invention can be deduced using methods well-known in the art. It will be understood that the sequence of an entire nucleic acid encoding an *H. pylori* polypeptide can be isolated and identified based on an ORF that encodes only a fragment of the cognate protein-coding region. This can be acheived, for example, by using the isolated nucleic acid encoding the ORF, or fragments thereof, to prime a polymerase chain reaction with genomic *H. pylori* DNA as template; this is followed by sequencing the amplified product.

The polypeptides of the invention can be isolated from wild-type or mutant *H. pylori* cells or from heterologous organisms or cells (including, but not limited to, bacteria, yeast, insect, plant and mammalian cells) into which an *H. pylori* nucleic acid has been introduced and expressed. In addition, the polypeptides can be part of recombinant fusion proteins.

*H. pylori* polypeptides of the invention can be chemically synthesized using commercially automated procedures such as those referenced herein.

IV. Identification of Nucleic Acids Encoding Vaccine Components and Targets for Agents Effective Against *H. pylori*

The disclosed *H. pylori* genome sequence includes segments that direct the synthesis of ribonucleic acids and polypeptides, as well as origins of replication, promoters, other types of regulatory sequences, and intergenic nucleic acids. The invention encompasses nucleic acids encoding immunogenic components of vaccines and targets for agents effective against *H. pylori*. Identification of said immunogenic components involved in the determination of the function of the disclosed sequences can be achieved using a variety of approaches. Non-limiting examples of these approaches are described briefly below.

Homology to known sequences: Computer-assisted comparison of the disclosed *H. pylori* sequences with previously reported sequences present in publicly available databases is useful for identifying functional *H. pylori* nucleic acid and polypeptide sequences. It will be understood that protein-coding sequences, for example, may be compared as a whole, and that a high degree of sequence homology between two proteins (such as, for example, >80-90%) at the amino acid level indicates that the two proteins also possess some degree of functional homology, such as, for example, among enzymes involved in metabolism, DNA synthesis, or cell wall synthesis, and proteins involved in transport, cell division, etc. In addition, many structural features of particular protein classes have been identified and correlate with specific consensus sequences, such as, for example, binding domains for nucleotides, DNA, metal ions, and other small molecules; sites for covalent modifications such as phosphorylation,

acylation, and the like; sites of protein:protein interactions, etc. These consensus sequences may be quite short and thus may represent only a fraction of the entire protein-coding sequence. Identification of such a feature in an *H. pylori* sequence is therefore useful in determining the function of the encoded protein and identifying useful targets of antibacterial drugs.

Of particular relevance to the present invention are structural features that are common to secretory, transmembrane, and surface proteins, including secretion signal peptides and hydrophobic transmembrane domains. *H. pylori* proteins identified as containing putative signal sequences and/or transmembrane domains are useful as immunogenic components of vaccines.

Identification of essential genes: Nucleic acids that encode proteins essential for growth or viability of *H. pylori* are preferred drug targets. *H. pylori* genes can be tested for their biological relevance to the organism by examining the effect of deleting and/or disrupting the genes, i.e., by so-called gene "knockout", using techniques known to those skilled in the relevant art. In this manner, essential genes may be identified.

Strain-specific sequences: Because of the evolutionary relationship between different *H. pylori* strains, it is believed that the presently disclosed *H. pylori* sequences are useful for identifying, and/or discriminating between, previously known and new *H. pylori* strains. It is believed that other *H. pylori* strains will exhibit at least 70% sequence homology with the presently disclosed sequence. Systematic and routine analyses of DNA sequences derived from samples containing *H. pylori* strains, and comparison with the present sequence allows for the identification of sequences that can be used to discriminate between strains, as well as those that are common to all *H. pylori* strains. In one embodiment, the invention provides nucleic acids, including probes, and peptide and polypeptide sequences that discriminate between different strains of *H. pylori*. Strain-specific components can also be identified functionally by their ability to elicit or react with antibodies that selectively recognize one or more *H. pylori* strains.

In another embodiment, the invention provides nucleic acids, including probes, and peptide and polypeptide sequences that are common to all *H. pylori* strains but are *not* found in other bacterial species.

Specific Example: Determination Of Candidate Protein Antigens For Antibody And Vaccine Development

The selection of candidate protein antigens for vaccine development can be derived from the nucleic acids encoding *H. pylori* polypeptides. First, the ORF's can be analyzed for homology to other known exported or membrane proteins and analyzed using the discriminant analysis described by Klein, et al. (Klein, P., Kanehsia, M., and

DeLisi, C. (1985) *Biochimica et Biophysica Acta* 815, 468-476) for predicting exported and membrane proteins.

Homology searches can be performed using the BLAST algorithm contained in the Wisconsin Sequence Analysis Package (Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711) to compare each predicted ORF amino acid sequence with all sequences found in the current GenBank, SWISS-PROT and PIR databases. BLAST searches for local alignments between the ORF and the databank sequences and reports a probability score which indicates the probability of finding this sequence by chance in the database. ORF's with significant homology (e.g. probabilities lower than $1x10^{-6}$ that the homology is only due to random chance) to membrane or exported proteins represent protein antigens for vaccine development. Possible functions can be provided to *H. pylori* genes based on sequence homology to genes cloned in other organisms.

Discriminant analysis (Klein, et al. supra) can be used to examine the ORF amino acid sequences. This algorithm uses the intrinsic information contained in the ORF amino acid sequence and compares it to information derived from the properties of known membrane and exported proteins. This comparison predicts which proteins will be exported, membrane associated or cytoplasmic. ORF amino acid sequences identified as exported or membrane associated by this algorithm are likely protein antigens for vaccine development.

Surface exposed outer membrane proteins are likely to represent the best antigens to provide a protective immune response against *H. pylori*. Among the algorithms that can be used to aid in prediction of these outer membrane proteins include the presence of an amphipathic beta-sheet region at their C-terminus. This region which has been detected in a large number of outer membrane proteins in Gram negative bacteria is often characterized by hydrophobic residues (Phe or Tyr) clustered at alternating positions from the C-terminus (e.g., see Figure 5, block F; Figure 7, block E). Importantly, these sequences have not been detected at the C-termini of periplasmic proteins, thus allowing preliminary distinction between these classes of proteins based on primary sequence data. This phenomenon has been reported previously by Struyve et al. (*J. Mol. Biol.* 218:141-148, 1991).

Also illustrated in Figure 5 are additional amino acid sequence motifs found in many outer membrane proteins of *H. pylori*. The amino acid sequence alignment in Figure 5 depicts portions of the sequence of five *H. pylori* proteins (depicted in the single letter amino acid code) labeled with their amino acid Sequence ID Numbers and shown N-terminal to C-terminal, left to right. Five or six distinct blocks (labeled A through E or F) of similar amino acid residues are found including the distinctive

hydrophobic residues (Phe or Tyr; F or Y according to the single letter code for amino acid residues) frequently found at positions near the C-terminus of outer membrane proteins. The presence of several shared motifs clearly establishes the similarity between members of this group of proteins.

Additional amino acid alignments for four outer membrane proteins isolated from *H. pylori* are depicted in Figure 6.

Outer membrane proteins isolated from *H. pylori* frequently share additional motifs as depicted for two proteins in Figure 7 which also share the C-terminal hydrophobic residues, and as depicted for two proteins in Figure 8 which do not share the C-terminal hydrophobic residue motif but share a different C-terminal motif.

One skilled in the art would know that these shared sequence motifs are highly significant and establish a similarity among this group of proteins.

Infrequently it is not possible to distinguish between multiple possible nucleotides at a given position in the nucleic acid sequence. In those cases the ambiguities are denoted by an extended alphabet as follows:

These are the official IUPAC-IUB single-letter base codes

| Code | Base Description | |
|---|---|---|
| G | Guanine | |
| A | Adenine | |
| T | Thymine | |
| C | Cytosine | |
| R | Purine | (A or G) |
| Y | Pyrimidine | (C or T or U) |
| M | Amino | (A or C) |
| K | Ketone | (G or T) |
| S | Strong interaction | (C or G) |
| W | Weak interaction | (A or T) |
| H | Not-G | (A or C or T) |
| B | Not-A | (C or G or T) |
| V | Not-T (not-U) | (A or C or G) |
| D | Not-C | (A or G or T) |
| N | Any | (A or C or G or T) |

The amino acid translations of this invention account for the ambiguity in the nucleic acid sequence by translating the ambiguous codon as the letter "X". In all cases,

the permissible amino acid residues at a position are clear from an examination of the nucleic acid sequence based on the standard genetic code.

V. Production of Fragments and Analogs of *H. pylori* Nucleic Acids and Polypeptides

Based on the discovery of the *H. pylori* gene products of the invention provided in the Sequence Lsiting, one skilled in the art can alter the disclosed structure (of *H. pylori* genes), e.g., by producing fragments or analogs, and test the newly produced structures for activity. Examples of techniques known to those skilled in the relevant art which allow the production and testing of fragments and analogs are discussed below. These, or analogous methods can be used to make and screen libraries of polypeptides, e.g., libraries of random peptides or libraries of fragments or analogs of cellular proteins for the ability to bind *H. pylori* polypeptides. Such screens are useful for the identification of inhibitors of *H. pylori*.

Generation of Fragments

Fragments of a protein can be produced in several ways, e.g., recombinantly, by proteolytic digestion, or by chemical synthesis. Internal or terminal fragments of a polypeptide can be generated by removing one or more nucleotides from one end (for a terminal fragment) or both ends (for an internal fragment) of a nucleic acid which encodes the polypeptide. Expression of the mutagenized DNA produces polypeptide fragments. Digestion with "end-nibbling" endonucleases can thus generate DNA's which encode an array of fragments. DNA's which encode fragments of a protein can also be generated by random shearing, restriction digestion or a combination of the above-discussed methods.

Fragments can also be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, peptides of the present invention may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or divided into overlapping fragments of a desired length.

Alteration of Nucleic Acids and Polypeptides: Random Methods

Amino acid sequence variants of a protein can be prepared by random mutagenesis of DNA which encodes a protein or a particular domain or region of a protein. Useful methods include PCR mutagenesis and saturation mutagenesis. A library of random amino acid sequence variants can also be generated by the synthesis of a set of degenerate oligonucleotide sequences. (Methods for screening proteins in a library of variants are elsewhere herein).

(A) PCR Mutagenesis

In PCR mutagenesis, reduced Taq polymerase fidelity is used to introduce random mutations into a cloned fragment of DNA (Leung et al., 1989, *Technique* 1:11-15). The DNA region to be mutagenized is amplified using the polymerase chain reaction (PCR) under conditions that reduce the fidelity of DNA synthesis by Taq DNA polymerase, e.g., by using a dGTP/dATP ratio of five and adding $Mn^{2+}$ to the PCR reaction. The pool of amplified DNA fragments are inserted into appropriate cloning vectors to provide random mutant libraries.

(B) Saturation Mutagenesis

Saturation mutagenesis allows for the rapid introduction of a large number of single base substitutions into cloned DNA fragments (Mayers et al., 1985, *Science* 229:242). This technique includes generation of mutations, e.g., by chemical treatment or irradiation of single-stranded DNA *in vitro*, and synthesis of a complimentary DNA strand. The mutation frequency can be modulated by modulating the severity of the treatment, and essentially all possible base substitutions can be obtained. Because this procedure does not involve a genetic selection for mutant fragments both neutral substitutions, as well as those that alter function, are obtained. The distribution of point mutations is not biased toward conserved sequence elements.

(C) Degenerate Oligonucleotides

A library of homologs can also be generated from a set of degenerate oligonucleotide sequences. Chemical synthesis of a degenerate sequences can be carried out in an automatic DNA synthesizer, and the synthetic genes then ligated into an appropriate expression vector. The synthesis of degenerate oligonucleotides is known in the art (see for example, Narang, SA (1983) *Tetrahedron* 39:3; Itakura et al. (1981) *Recombinant DNA, Proc 3rd Cleveland Sympos. Macromolecules*, ed. AG Walton, Amsterdam: Elsevier pp273-289; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477. Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al. (1990) *Science* 249:386-390; Roberts et al. (1992) *PNAS* 89:2429-2433; Devlin et al. (1990) *Science* 249: 404-406; Cwirla et al. (1990) *PNAS* 87: 6378-6382; as well as U.S. Patents Nos. 5,223,409, 5,198,346, and 5,096,815).

Alteration of Nucleic Acids and Polypeptides: Methods for Directed Mutagenesis

Non-random or directed, mutagenesis techniques can be used to provide specific sequences or mutations in specific regions. These techniques can be used to create variants which include, e.g., deletions, insertions, or substitutions, of residues of the known amino acid sequence of a protein. The sites for mutation can be modified individually or in series, e.g., by (1) substituting first with conserved amino acids and then with more radical choices depending upon results achieved, (2) deleting the target residue, or (3) inserting residues of the same or a different class adjacent to the located site, or combinations of options 1-3.

(A) Alanine Scanning Mutagenesis

Alanine scanning mutagenesis is a useful method for identification of certain residues or regions of the desired protein that are preferred locations or domains for mutagenesis, Cunningham and Wells (*Science* 244:1081-1085, 1989). In alanine scanning, a residue or group of target residues are identified (e.g., charged residues such as Arg, Asp, His, Lys, and Glu) and replaced by a neutral or negatively charged amino acid (most preferably alanine or polyalanine). Replacement of an amino acid can affect the interaction of the amino acids with the surrounding aqueous environment in or outside the cell. Those domains demonstrating functional sensitivity to the substitutions are then refined by introducing further or other variants at or for the sites of substitution. Thus, while the site for introducing an amino acid sequence variation is predetermined, the nature of the mutation per se need not be predetermined. For example, to optimize the performance of a mutation at a given site, alanine scanning or random mutagenesis may be conducted at the target codon or region and the expressed desired protein subunit variants are screened for the optimal combination of desired activity.

(B) Oligonucleotide-Mediated Mutagenesis

Oligonucleotide-mediated mutagenesis is a useful method for preparing substitution, deletion, and insertion variants of DNA, see, e.g., Adelman et al., (*DNA* 2:183, 1983). Briefly, the desired DNA is altered by hybridizing an oligonucleotide encoding a mutation to a DNA template, where the template is the single-stranded form of a plasmid or bacteriophage containing the unaltered or native DNA sequence of the desired protein. After hybridization, a DNA polymerase is used to synthesize an entire second complementary strand of the template that will thus incorporate the oligonucleotide primer, and will code for the selected alteration in the desired protein DNA. Generally, oligonucleotides of at least 25 nucleotides in length are used. An optimal oligonucleotide will have 12 to 15 nucleotides that are completely

complementary to the template on either side of the nucleotide(s) coding for the mutation. This ensures that the oligonucleotide will hybridize properly to the single-stranded DNA template molecule. The oligonucleotides are readily synthesized using techniques known in the art such as that described by Crea et al. (*Proc. Natl. Acad. Sci.* USA, 75: 5765[1978]).

(C) Cassette Mutagenesis

Another method for preparing variants, cassette mutagenesis, is based on the technique described by Wells et al. (*Gene*, 34:315[1985]). The starting material is a plasmid (or other vector) which includes the protein subunit DNA to be mutated. The codon(s) in the protein subunit DNA to be mutated are identified. There must be a unique restriction endonuclease site on each side of the identified mutation site(s). If no such restriction sites exist, they may be generated using the above-described oligonucleotide-mediated mutagenesis method to introduce them at appropriate locations in the desired protein subunit DNA. After the restriction sites have been introduced into the plasmid, the plasmid is cut at these sites to linearize it. A double-stranded oligonucleotide encoding the sequence of the DNA between the restriction sites but containing the desired mutation(s) is synthesized using standard procedures. The two strands are synthesized separately and then hybridized together using standard techniques. This double-stranded oligonucleotide is referred to as the cassette. This cassette is designed to have 3' and 5' ends that are comparable with the ends of the linearized plasmid, such that it can be directly ligated to the plasmid. This plasmid now contains the mutated desired protein subunit DNA sequence.

(D) Combinatorial Mutagenesis

Combinatorial mutagenesis can also be used to generate mutants (Ladner et al., WO 88/06630). In this method, the amino acid sequences for a group of homologs or other related proteins are aligned, preferably to promote the highest homology possible. All of the amino acids which appear at a given position of the aligned sequences can be selected to create a degenerate set of combinatorial sequences. The variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level, and is encoded by a variegated gene library. For example, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential sequences are expressible as individual peptides, or alternatively, as a set of larger fusion proteins containing the set of degenerate sequences.

Other Modifications of *H. pylori* Nucleic Acids and Polypeptides

It is possible to modify the structure of an *H. pylori* polypeptide for such purposes as increasing solubility, enhancing stability (e.g., shelf life *ex vivo* and resistance to proteolytic degradation *in vivo*). A modified *H. pylori* protein or peptide can be produced in which the amino acid sequence has been altered, such as by amino acid substitution, deletion, or addition as described herein.

An *H. pylori* peptide can also be modified by substitution of cysteine residues preferably with alanine, serine, threonine, leucine or glutamic acid residues to minimize dimerization via disulfide linkages. In addition, amino acid side chains of fragments of the protein of the invention can be chemically modified. Another modification is cyclization of the peptide.

In order to enhance stability and/or reactivity, an *H. pylori* polypeptide can be modified to incorporate one or more polymorphisms in the amino acid sequence of the protein resulting from any natural allelic variation. Additionally, D-amino acids, non-natural amino acids, or non-amino acid analogs can be substituted or added to produce a modified protein within the scope of this invention. Furthermore, an *H. pylori* polypeptide can be modified using polyethylene glycol (PEG) according to the method of A. Sehon and co-workers (Wie et al., supra) to produce a protein conjugated with PEG. In addition, PEG can be added during chemical synthesis of the protein. Other modifications of *H. pylori* proteins include reduction/alkylation (Tarr, *Methods of Protein Microcharacterization*, J. E. Silver ed., Humana Press, Clifton NJ 155-194 (1986)); acylation (Tarr, supra); chemical coupling to an appropriate carrier (Mishell and Shiigi, eds, *Selected Methods in Cellular Immunology*, WH Freeman, San Francisco, CA (1980), U.S. Patent 4,939,239; or mild formalin treatment (Marsh, (1971) *Int. Arch. of Allergy and Appl. Immunol.*, 41: 199 - 215).

To facilitate purification and potentially increase solubility of an *H. pylori* protein or peptide, it is possible to add an amino acid fusion moiety to the peptide backbone. For example, hexa-histidine can be added to the protein for purification by immobilized metal ion affinity chromatography (Hochuli, E. et al., (1988) *Bio/Technology*, 6: 1321 - 1325). In addition, to facilitate isolation of peptides free of irrelevant sequences, specific endoprotease cleavage sites can be introduced between the sequences of the fusion moiety and the peptide.

To potentially aid proper antigen processing of epitopes within an *H. pylori* polypeptide, canonical protease sensitive sites can be engineered between regions, each comprising at least one epitope via recombinant or synthetic methods. For example, charged amino acid pairs, such as KK or RR, can be introduced between regions within a protein or fragment during recombinant construction thereof. The resulting peptide

can be rendered sensitive to cleavage by cathepsin and/or other trypsin-like enzymes which would generate portions of the protein containing one or more epitopes. In addition, such charged amino acid residues can result in an increase in the solubility of the peptide.

Primary Methods for Screening Polypeptides and Analogs

Various techniques are known in the art for screening generated mutant gene products. Techniques for screening large gene libraries often include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the genes under conditions in which detection of a desired activity, e.g., in this case, binding to *H. pylori* polypeptide or an interacting protein, facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the techniques described below is amenable to high through-put analysis for screening large numbers of sequences created, e.g., by random mutagenesis techniques.

(A) Two Hybrid Systems

Two hybrid assays such as the system described above (as with the other screening methods described herein), can be used to identify polypeptides, e.g., fragments or analogs of a naturally-occurring *H. pylori* polypeptide, e.g., of cellular proteins, or of randomly generated polypeptides which bind to an *H. pylori* protein. (The *H. pylori* domain is used as the bait protein and the library of variants are expressed as fish fusion proteins.) In an analogous fashion, a two hybrid assay (as with the other screening methods described herein), can be used to find polypeptides which bind a *H. pylori* polypeptide.

(B) Display Libraries

In one approach to screening assays, the candidate peptides are displayed on the surface of a cell or viral particle, and the ability of particular cells or viral particles to bind an appropriate receptor protein via the displayed product is detected in a "panning assay". For example, the gene library can be cloned into the gene for a surface membrane protein of a bacterial cell, and the resulting fusion protein detected by panning (Ladner et al., WO 88/06630; Fuchs et al. (1991) *Bio/Technology* 9:1370-1371; and Goward et al. (1992) *TIBS* 18:136-140). In a similar fashion, a detectably labeled ligand can be used to score for potentially functional peptide homologs. Fluorescently labeled ligands, e.g., receptors, can be used to detect homologs which retain ligand-binding activity. The use of fluorescently labeled ligands, allows cells to be visually

inspected and separated under a fluorescence microscope, or, where the morphology of the cell permits, to be separated by a fluorescence-activated cell sorter.

A gene library can be expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences can be expressed on the surface of infectious phage, thereby conferring two significant benefits. First, since these phage can be applied to affinity matrices at concentrations well over $10^{13}$ phage per milliliter, a large number of phage can be screened at one time. Second, since each infectious phage displays a gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage can be amplified by another round of infection. The group of almost identical *E. coli* filamentous phages M13, fd., and f1 are most often used in phage display libraries. Either of the phage gIII or gVIII coat proteins can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle. Foreign epitopes can be expressed at the $NH_2$-terminal end of pIII and phage bearing such epitopes recovered from a large excess of phage lacking this epitope (Ladner et al. PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al. (1992) *J. Biol. Chem.* 267:16007-16010; Griffiths et al. (1993) *EMBO J* 12:725-734; Clackson et al. (1991) *Nature* 352:624-628; and Barbas et al. (1992) *PNAS* 89:4457-4461).

A common approach uses the maltose receptor of *E. coli* (the outer membrane protein, LamB) as a peptide fusion partner (Charbit et al. (1986) *EMBO* 5, 3029-3037). Oligonucleotides have been inserted into plasmids encoding the LamB gene to produce peptides fused into one of the extracellular loops of the protein. These peptides are available for binding to ligands, e.g., to antibodies, and can elicit an immune response when the cells are administered to animals. Other cell surface proteins, e.g., OmpA (Schorr et al. (1991) *Vaccines 91*, pp. 387-392), PhoE (Agterberg, et al. (1990) *Gene* 88, 37-45), and PAL (Fuchs et al. (1991) *Bio/Tech* 9, 1369-1372), as well as large bacterial surface structures have served as vehicles for peptide display. Peptides can be fused to pilin, a protein which polymerizes to form the pilus-a conduit for interbacterial exchange of genetic information (Thiry et al. (1989) *Appl. Environ. Microbiol.* 55, 984-993). Because of its role in interacting with other cells, the pilus provides a useful support for the presentation of peptides to the extracellular environment. Another large surface structure used for peptide display is the bacterial motive organ, the flagellum. Fusion of peptides to the subunit protein flagellin offers a dense array of many peptide copies on the host cells (Kuwajima et al. (1988) *Bio/Tech.* 6, 1080-1083). Surface proteins of other bacterial species have also served as peptide fusion partners. Examples include the *Staphylococcus* protein A and the outer membrane IgA protease of *Neisseria* (Hansson

et al. (1992) *J. Bacteriol.* 174, 4239-4245 and Klauser et al. (1990) *EMBO J.* 9, 1991-1999).

In the filamentous phage systems and the LamB system described above, the physical link between the peptide and its encoding DNA occurs by the containment of the DNA within a particle (cell or phage) that carries the peptide on its surface. Capturing the peptide captures the particle and the DNA within. An alternative scheme uses the DNA-binding protein LacI to form a link between peptide and DNA (Cull *et al.* (1992) *PNAS USA* 89:1865-1869). This system uses a plasmid containing the LacI gene with an oligonucleotide cloning site at its 3'-end. Under the controlled induction by arabinose, a LacI-peptide fusion protein is produced. This fusion retains the natural ability of LacI to bind to a short DNA sequence known as LacO operator (LacO). By installing two copies of LacO on the expression plasmid, the LacI-peptide fusion binds tightly to the plasmid that encoded it. Because the plasmids in each cell contain only a single oligonucleotide sequence and each cell expresses only a single peptide sequence, the peptides become specifically and stably associated with the DNA sequence that directed its synthesis. The cells of the library are gently lysed and the peptide-DNA complexes are exposed to a matrix of immobilized receptor to recover the complexes containing active peptides. The associated plasmid DNA is then reintroduced into cells for amplification and DNA sequencing to determine the identity of the peptide ligands. As a demonstration of the practical utility of the method, a large random library of dodecapeptides was made and selected on a monoclonal antibody raised against the opioid peptide dynorphin B. A cohort of peptides was recovered, all related by a consensus sequence corresponding to a six-residue portion of dynorphin B. (Cull et al. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89-1869)

This scheme, sometimes referred to as peptides-on-plasmids, differs in two important ways from the phage display methods. First, the peptides are attached to the C-terminus of the fusion protein, resulting in the display of the library members as peptides having free carboxy termini. Both of the filamentous phage coat proteins, pIII and pVIII, are anchored to the phage through their C-termini, and the guest peptides are placed into the outward-extending N-terminal domains. In some designs, the phage-displayed peptides are presented right at the amino terminus of the fusion protein. (Cwirla, et al. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 6378-6382) A second difference is the set of biological biases affecting the population of peptides actually present in the libraries. The LacI fusion molecules are confined to the cytoplasm of the host cells. The phage coat fusions are exposed briefly to the cytoplasm during translation but are rapidly secreted through the inner membrane into the periplasmic compartment, remaining anchored in the membrane by their C-terminal hydrophobic domains, with the

N-termini, containing the peptides, protruding into the periplasm while awaiting assembly into phage particles. The peptides in the LacI and phage libraries may differ significantly as a result of their exposure to different proteolytic activities. The phage coat proteins require transport across the inner membrane and signal peptidase processing as a prelude to incorporation into phage. Certain peptides exert a deleterious effect on these processes and are underrepresented in the libraries (Gallop et al. (1994) *J. Med. Chem.* 37(9):1233-1251). These particular biases are not a factor in the LacI display system.

The number of small peptides available in recombinant random libraries is enormous. Libraries of $10^7$-$10^9$ independent clones are routinely prepared. Libraries as large as $10^{11}$ recombinants have been created, but this size approaches the practical limit for clone libraries. This limitation in library size occurs at the step of transforming the DNA containing randomized segments into the host bacterial cells. To circumvent this limitation, an *in vitro* system based on the display of nascent peptides in polysome complexes has recently been developed. This display library method has the potential of producing libraries 3-6 orders of magnitude larger than the currently available phage/phagemid or plasmid libraries. Furthermore, the construction of the libraries, expression of the peptides, and screening, is done in an entirely cell-free format.

In one application of this method (Gallop et al. (1994) *J. Med. Chem.* 37(9):1233-1251), a molecular DNA library encoding $10^{12}$ decapeptides was constructed and the library expressed in an *E. coli* S30 *in vitro* coupled transcription/translation system. Conditions were chosen to stall the ribosomes on the mRNA, causing the accumulation of a substantial proportion of the RNA in polysomes and yielding complexes containing nascent peptides still linked to their encoding RNA. The polysomes are sufficiently robust to be affinity purified on immobilized receptors in much the same way as the more conventional recombinant peptide display libraries are screened. RNA from the bound complexes is recovered, converted to cDNA, and amplified by PCR to produce a template for the next round of synthesis and screening. The polysome display method can be coupled to the phage display system. Following several rounds of screening, cDNA from the enriched pool of polysomes was cloned into a phagemid vector. This vector serves as both a peptide expression vector, displaying peptides fused to the coat proteins, and as a DNA sequencing vector for peptide identification. By expressing the polysome-derived peptides on phage, one can either continue the affinity selection procedure in this format or assay the peptides on individual clones for binding activity in a phage ELISA, or for binding specificity in a completion phage ELISA (Barret, et al. (1992) *Anal. Biochem* 204,357-364). To

identify the sequences of the active peptides one sequences the DNA produced by the phagemid host.

Secondary Screening of Polypeptides and Analogs

The high through-put assays described above can be followed by secondary screens in order to identify further biological activities which will, e.g., allow one skilled in the art to differentiate agonists from antagonists. The type of a secondary screen used will depend on the desired activity that needs to be tested. For example, an assay can be developed in which the ability to inhibit an interaction between a protein of interest and its respective ligand can be used to identify antagonists from a group of peptide fragments isolated though one of the primary screens described above.

Therefore, methods for generating fragments and analogs and testing them for activity are known in the art. Once the core sequence of interest is identified, it is routine for one skilled in the art to obtain analogs and fragments.

Peptide Mimetics of *H. pylori* Polypeptides

The invention also provides for reduction of the protein binding domains of the subject *H. pylori* polypeptides to generate mimetics, e.g. peptide or non-peptide agents. The peptide mimetics are able to disrupt binding of a polypeptide to its counter ligand, e.g., in the case of an *H. pylori* polypeptide binding to a naturally occurring ligand. The critical residues of a subject *H. pylori* polypeptide which are involved in molecular recognition of a polypeptide can be determined and used to generate *H. pylori*-derived peptidomimetics which competitively or noncompetitively inhibit binding of the *H. pylori* polypeptide with an interacting polypeptide (see, for example, European patent applications EP-412,762A and EP-B31,080A).

For example, scanning mutagenesis can be used to map the amino acid residues of a particular *H. pylori* polypeptide involved in binding an interacting polypeptide, peptidomimetic compounds (e.g. diazepine or isoquinoline derivatives) can be generated which mimic those residues in binding to an interacting polypeptide, and which therefore can inhibit binding of an *H. pylori* polypeptide to an interacting polypeptide and thereby interfere with the function of *H. pylori* polypeptide. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gama lactam rings (Garvey et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-

methylene pseudopeptides (Ewenson et al. (1986) *J Med Chem* 29:295; and Ewenson et al. in *Peptides: Structure and Function* (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, IL, 1985), β-turn dipeptide cores (Nagai et al. (1985) *Tetrahedron Lett* 26:647; and Sato et al. (1986) *J Chem Soc Perkin Trans* 1:1231), and β-aminoalcohols (Gordon et al. (1985) *Biochem Biophys Res Commun*126:419; and Dann et al. (1986) *Biochem Biophys Res Commun* 134:71).

VI. Vaccine Formulations for *H. pylori* Nucleic Acids and Polypeptides

This invention also features vaccine compositions or formulations (used interchangeably herein) for protection against infection by *H. pylori* or for treatment of *H. pylori* infection. As used herein, the term "treatment of *H. pylori* infection" refers to therapeutic treatment of an existing or established *H. pylori* infection. The terms "protection against *H. pylori* infection" or "prophylactic treatment" refer to the use of *H. pylori* vaccine formulation for reducing the risk of or preventing an infection in a subject at risk for *H. pylori* infection. In one embodiment, the vaccine compositions contain one or more immunogenic components, such as a surface protein, from *H. pylori*, or portion thereof, and a pharmaceutically acceptable carrier. For example, in one embodiment, the vaccine formulations of the invention contain at least one or combination of *H. pylori* polypeptides or fragments thereof, from same or different *H. pylori* antigens. Nucleic acids and *H. pylori* polypeptides for use in the vaccine formulations of the invention include the nucleic acids and polypeptides set forth in the Sequence Listing, preferably those *H. pylori* nucleic acids that encode surface proteins and surface proteins or fragments thereof. For example, a preferred nucleic acid and *H. pylori* polypeptide for use in a vaccine composition of the invention is selected from the group of nucleic acids which encode cell envelope proteins and *H. pylori* cell envelope proteins as set forth in Table 1. However, any nucleic acid encoding an immunogenic *H. pylori* protein and *H. pylori* polypetide, or portion thereof, can be used in the present invention. These vaccines have therapeutic and/or prophylactic utilities.

One aspect of the invention provides a vaccine composition for protection against infection by *H. pylori* which contains at least one immunogenic fragment of an *H. pylori* protein and a pharmaceutically acceptable carrier. Preferred fragments include peptides of at least about 10 amino acid residues in length, preferably about 10-20 amino acid residues in length, and more preferably about 12-16 amino acid residues in length.

Immunogenic components of the invention can be obtained, for example, by screening polypeptides recombinantly produced from the corresponding fragment of the nucleic acid encoding the full-length *H. pylori* protein. In addition, fragments can be

chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry.

In one embodiment, immunogenic components are identified by the ability of the peptide to stimulate T cells. Peptides which stimulate T cells, as determined by, for example, T cell proliferation or cytokine secretion are defined herein as comprising at least one T cell epitope. T cell epitopes are believed to be involved in initiation and perpetuation of the immune response to the protein allergen which is responsible for the clinical symptoms of allergy. These T cell epitopes are thought to trigger early events at the level of the T helper cell by binding to an appropriate HLA molecule on the surface of an antigen presenting cell, thereby stimulating the T cell subpopulation with the relevant T cell receptor for the epitope. These events lead to T cell proliferation, lymphokine secretion, local inflammatory reactions, recruitment of additional immune cells to the site of antigen/T cell interaction, and activation of the B cell cascade, leading to the production of antibodies. A T cell epitope is the basic element, or smallest unit of recognition by a T cell receptor, where the epitope comprises amino acids essential to receptor recognition (e.g., approximately 6 or 7 amino acid residues). Amino acid sequences which mimic those of the T cell epitopes are within the scope of this invention.

In another embodiment, immunogenic components of the invention are identified through genomic vaccination. The basic protocol is based on the idea that expression libraries consisting of all or parts of a pathogen genome, e.g., an *H. pylori* genome, can confer protection when used to genetically immunize a host. This expression library immunization (ELI) is analogous to expression cloning and involves reducing a genomic expression library of a pathogen, e.g., *H. pylori*, into plasmids that can act as genetic vaccines. The plasmids can also be designed to encode genetic adjuvants which can dramatically stimulate the humoral response. These genetic adjuvants can be introduced at remote sites and act as well extracelluraly as intracellularly.

This is a new approach to vaccine production that has many of the advantages of live/attenuated pathogens but no risk of infection. An expression library of pathogen DNA is used to immunize a host thereby producing the effects of antigen presentation of a live vaccine without the risk. For example, in the present invention, random fragments from the *H. pylori* genome or from cosmid or plasmid clones, as well as PCR products from genes identified by genomic sequencing, can be used to immunize a host. The feasibility of this approach has been demonstrated with *Mycoplasma pulmonis* (Barry et al., *Nature* 377:632-635, 1995), where even partial expression libraries of *Mycoplasma pulmonis*, a natural pathogen in rodents, provided protection against challenge from the pathogen.

ELI is a technique that allows for production of a non-infectious multipartite vaccine, even when little is known about pathogen's biology, because ELI uses the immune system to screen candidate genes. Once isolated, these genes can be used as genetic vaccines or for development of recombinant protein vaccines. Thus, ELI allows for production of vaccines in a systematic, largely mechanized fashion.

Screening immunogenic components can be accomplished using one or more of several different assays. For example, *in vitro*, peptide T cell stimulatory activity is assayed by contacting a peptide known or suspected of being immunogenic with an antigen presenting cell which presents appropriate MHC molecules in a T cell culture. Presentation of an immunogenic *H. pylori* peptide in association with appropriate MHC molecules to T cells in conjunction with the necessary costimulation has the effect of transmitting a signal to the T cell that induces the production of increased levels of cytokines, particularly of interleukin-2 and interleukin-4. The culture supernatant can be obtained and assayed for interleukin-2 or other known cytokines. For example, any one of several conventional assays for interleukin-2 can be employed, such as the assay described in *Proc. Natl. Acad. Sci USA*, 86: 1333 (1989) the pertinent portions of which are incorporated herein by reference. A kit for an assay for the production of interferon is also available from Genzyme Corporation (Cambridge, MA).

Alternatively, a common assay for T cell proliferation entails measuring tritiated thymidine incorporation. The proliferation of T cells can be measured *in vitro* by determining the amount of $^{3}H$-labeled thymidine incorporated into the replicating DNA of cultured cells. Therefore, the rate of DNA synthesis and, in turn, the rate of cell division can be quantified.

Vaccine compositions or formulations of the invention containing one or more immunogenic components (e.g., *H. pylori* polypeptide or fragment thereof or nucleic acid encoding an *H. pylori* polypeptide or fragment thereof) preferably include a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable pharmaceutically acceptable carriers include, for example, one or more of water, saline, phosphate buffered saline, dextrose, glycerol, ethanol and the like, as well as combinations thereof. Pharmaceutically acceptable carriers may further comprise minor amounts of auxiliary substances such as wetting or emulsifying agents, preservatives or buffers, which enhance the shelf life or effectiveness of the *H. pylori* nucleic acid or polypeptide. For vaccine formulations of the invention containing *H. pylori* polypeptides, the polypeptide is preferably coadministered with a suitable adjuvant and/or a delivery system described herein.

It will be apparent to those of skill in the art that the therapeutically effective amount of DNA or protein of this invention will depend, *inter alia*, upon the administration schedule, the unit dose of an *H. pylori* nucleic acid or polypeptide administered, whether the protein or nucleic acid is administered in combination with other therapeutic agents, the immune status and health of the patient, and the therapeutic activity of the particular protein or nucleic acid.

Vaccine formulations are conventionally administered parenterally, e.g., by injection, either subcutaneously or intramuscularly. Methods for intramuscular immunization are described by Wolff et al. (1990) *Science* 247: 1465-1468 and by Sedegah et al. (1994) *Immunology* 91: 9866-9870. Other modes of administration include oral and pulmonary formulations, suppositories, and transdermal applications. Oral immunization is preferred over parenteral methods for inducing protection against infection by *H. pylori*. Czinn et. al. (1993) *Vaccine* 11: 637-642. Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like.

In one embodiment, the vaccine formulation includes, as a pharmaceutically acceptable carrier, an adjuvant. Examples of the suitable adjuvants for use in the vaccine formulations of the invention include, but are not limited, to aluminum hydroxide; N-acetyl-muramyl--L-threonyl-D-isoglutamine (thr-MDP); N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP); N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphos-phoryloxy)-ethylamine (CGP 19835A, referred to a MTP-PE); RIBI, which contains three components from bacteria; monophosphoryl lipid A; trehalose dimycoloate; cell wall skeleton (MPL + TDM + CWS) in a 2% squalene/Tween 80 emulsion; and cholera toxin. Others which may be used are non-toxic derivatives of cholera toxin, including its B subunit, and/or conjugates or genetically engineered fusions of the *H. pylori* polypeptide with cholera toxin or its B subunit, procholeragenoid, fungal polysaccharides, including schizophyllan, muramyl dipeptide, muramyl dipeptide derivatives, phorbol esters, labile toxin of *E. coli*, non-*H. pylori* bacterial lysates, block polymers or saponins.

In another embodiment, the vaccine formulation includes, as a pharmaceutically acceptable carrier, a delivery system. Suitable delivery systems for use in the vaccine formulations of the invention include biodegradable microcapsules or immuno-stimulating complexes (ISCOMs), cochleates, or liposomes, genetically engineered attenuated live vectors such as viruses or bacteria, and recombinant (chimeric) virus-like

particles, e.g., bluetongue. In another embodiment of the invention, the vaccine formulation includes both a delivery system and an adjuvant.

Delivery systems in humans may include enteric release capsules protecting the antigen from the acidic environment of the stomach, and including *H. pylori* polypeptide in an insoluble form as fusion proteins. Suitable carriers for the vaccines of the invention are enteric coated capsules and polylactide-glycolide microspheres. Suitable diluents are 0.2 N NaHCO3 and/or saline.

Vaccines of the invention can be administered as a primary prophylactic agent in adults or in children, as a secondary prevention, after successful eradication of *H. pylori* in an infected host, or as a therapeutic agent in the aim to induce an immune response in a susceptible host to prevent infection by *H. pylori*. The vaccines of the invention are administered in amounts readily determined by persons of ordinary skill in the art. Thus, for adults a suitable dosage will be in the range of 10 μg to 10 g, preferably 10 μg to 100 mg, for example 50 μg to 50 mg. A suitable dosage for adults will also be in the range of 5 μg to 500 mg. Similar dosage ranges will be applicable for children.

The amount of adjuvant employed will depend on the type of adjuvant used. For example, when the mucosal adjuvant is cholera toxin, it is suitably used in an amount of 5 μg to 50 μg, for example 10 μg to 35 μg. When used in the form of microcapsules, the amount used will depend on the amount employed in the matrix of the microcapsule to achieve the desired dosage. The determination of this amount is within the skill of a person of ordinary skill in the art.

Those skilled in the art will recognize that the optimal dose may be more or less depending upon the patient's body weight, disease, the route of administration, and other factors. Those skilled in the art will also recognize that appropriate dosage levels can be obtained based on results with known oral vaccines such as, for example, a vaccine based on an *E. coli* lysate (6 mg dose daily up to total of 540 mg) and with an enterotoxigenic *E. coli* purified antigen (4 doses of 1 mg) (Schulman et al., *J. Urol.* 150:917-921 (1993)); Boedecker et al., *American Gastroenterological Assoc.* 999:A-222 (1993)). The number of doses will depend upon the disease, the formulation, and efficacy data from clinical trials. Without intending any limitation as to the course of treatment, the treatment can be administered over 3 to 8 doses for a primary immunization schedule over 1 month (Boedeker, *American Gastroenterological Assoc.* 888:A-222 (1993)).

In a preferred embodiment, a vaccine composition of the invention can be based on a killed whole *E. coli* preparation with an immunogenic fragment of an *H. pylori* protein of the invention expressed on its surface or it can be based on an *E. coli* lysate, wherein the killed *E. coli* acts as a carrier or an adjuvant.

It will be apparent to those skilled in the art that some of the vaccine compositions of the invention are useful only for preventing *H. pylori* infection, some are useful only for treating *H. pylori* infection, and some are useful for both preventing and treating *H. pylori* infection. In a preferred embodiment, the vaccine composition of the invention provides protection against *H. pylori* infection by stimulating humoral and/or cell-mediated immunity against *H. pylori*. It should be understood that amelioration of any of the symptoms of *H. pylori* infection is a desirable clinical goal, including a lessening of the dosage of medication used to treat *H. pylori*-caused disease, or an increase in the production of antibodies in the serum or mucous of patients.

VII. Antibodies Reactive With *H. pylori* Polypeptides

The invention also includes antibodies specifically reactive with the subject *H. pylori* polypeptide. Anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols (See, for example, *Antibodies: A Laboratory Manual* ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal such as a mouse, a hamster or rabbit can be immunized with an immunogenic form of the peptide. Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic portion of the subject *H. pylori* polypeptide can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies.

In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of the *H. pylori* polypeptides of the invention, e.g. antigenic determinants of a polypeptide of the invention contained in the Sequence Listing, or a closely related human or non-human mammalian homolog (e.g., 90% homologous, more preferably at least 95% homologous). In yet a further preferred embodiment of the invention, the anti-*H. pylori* antibodies do not substantially cross react (i.e., react specifically) with a protein which is for example, less than 80% percent homologous to a sequence of the invention contained in the Sequence Listing. By "not substantially cross react", it is meant that the antibody has a binding affinity for a non-homologous protein which is less than 10 percent, more preferably less than 5 percent, and even more preferably less than 1 percent, of the binding affinity for a protein of the invention contained in the Sequence Listing. In a most preferred embodiment, there is no crossreactivity between bacterial and mammalian antigens.

The term antibody as used herein is intended to include fragments thereof which are also specifically reactive with *H. pylori* polypeptides. Antibodies can be fragmented

using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, $F(ab')_2$ fragments can be generated by treating antibody with pepsin. The resulting $F(ab')_2$ fragment can be treated to reduce disulfide bridges to produce Fab' fragments. The antibody of the invention is further intended to include bispecific and chimeric molecules having an anti-*H. pylori* portion.

Both monoclonal and polyclonal antibodies (Ab) directed against *H. pylori* polypeptides or *H. pylori* polypeptide variants, and antibody fragments such as Fab` and $F(ab`)_2$, can be used to block the action of *H. pylori* polypeptide and allow the study of the role of a particular *H. pylori* polypeptide of the invention in aberrant or unwanted intracellular signaling, as well as the normal cellular function of the *H. pylori* and by microinjection of anti-*H. pylori* polypeptide antibodies of the present invention.

Antibodies which specifically bind *H. pylori* epitopes can also be used in immunohistochemical staining of tissue samples in order to evaluate the abundance and pattern of expression of *H. pylori* antigens. Anti *H. pylori* polypeptide antibodies can be used diagnostically in immuno-precipitation and immuno-blotting to detect and evaluate *H. pylori* levels in tissue or bodily fluid as part of a clinical testing procedure. Likewise, the ability to monitor *H. pylori* polypeptide levels in an individual can allow determination of the efficacy of a given treatment regimen for an individual afflicted with such a disorder. The level of an *H. pylori* polypeptide can be measured in cells found in bodily fluid, such as in urine samples or can be measured in tissue, such as produced by gastric biopsy. Diagnostic assays using anti-*H. pylori* antibodies can include, for example, immunoassays designed to aid in early diagnosis of *H. pylori* infections. The present invention can also be used as a method of detecting antibodies contained in samples from individuals infected by this bacterium using specific *H. pylori* antigens.

Another application of anti-*H. pylori* polypeptide antibodies of the invention is in the immunological screening of cDNA libraries constructed in expression vectors such as λgt11, λgt18-23, λZAP, and λORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance, λgt11 will produce fusion proteins whose amino termini consist of ß-galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of a subject *H. pylori* polypeptide can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with anti-*H. pylori* polypeptide antibodies. Phage, scored by this assay, can then be isolated from the infected plate. Thus, the presence of *H. pylori* gene

homologs can be detected and cloned from other species, and alternate isoforms (including splicing variants) can be detected and cloned.

VIII. Kits Containing Nucleic Acids, Polypeptides or Antibodies of the Invention

The nucleic acid, polypeptides and antibodies of the invention can be combined with other reagents and articles to form kits. Kits for diagnostic purposes typically comprise the nucleic acid, polypeptides or antibodies in vials or other suitable vessels. Kits typically comprise other reagents for performing hybridization reactions, polymerase chain reactions (PCR), or for reconstitution of lyophilized components, such as aqueous media, salts, buffers, and the like. Kits may also comprise reagents for sample processing such as detergents, chaotropic salts and the like. Kits may also comprise immobilization means such as particles, supports, wells, dipsticks and the like. Kits may also comprise labeling means such as dyes, developing reagents, radioisotopes, fluorescent agents, luminescent or chemiluminescent agents, enzymes, intercalating agents and the like. With the nucleic acid and amino acid sequence information provided herein, individuals skilled in art can readily assemble kits to serve their particular purpose. Kits further can include instructions for use.

IX. Drug Screening Assays Using *H. pylori* Polypeptides

By making available purified and recombinant *H. pylori* polypeptides, the present invention provides assays which can be used to screen for drugs which are either agonists or antagonists of the normal cellular function, in this case, of the subject *H. pylori* polypeptides, or of their role in intracellular signaling. Such inhibitors or potentiators may be useful as new therapeutic agents to combat *H. pylori* infections in humans. A variety of assay formats will suffice and, in light of the present inventions, will be comprehended by the skilled artisan.

In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays which are performed in cell-free systems, such as may be derived with purified or semi-purified proteins, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the *in vitro* system, the assay instead being focused primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with other proteins or change in enzymatic properties of the molecular target. Accordingly, in an exemplary screening assay of the present

invention, the compound of interest is contacted with an isolated and purified *H. pylori* polypeptide.

Screening assays can be constructed *in vitro* with a purified *H. pylori* polypeptide or fragment thereof, such as an *H. pylori* polypeptide having enzymatic activity, such that the activity of the polypeptide produces a detectable reaction product. The efficacy of the compound can be assessed by generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. Suitable products include those with distinctive absorption, fluorescence, or chemi-luminescence properties, for example, because detection may be easily automated. A variety of synthetic or naturally occurring compounds can be tested in the assay to identify those which inhibit or potentiate the activity of the *H. pylori* polypeptide. Some of these active compounds may directly, or with chemical alterations to promote membrane permeability or solubility, also inhibit or potentiate the same activity (e.g., enzymatic activity) in whole, live *H. pylori* cells.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references and published patent applications cited throughout this application are hereby incorporated by reference.

## EXEMPLIFICATION

I. Cloning and Sequencing of *H. pylori* DNA

*H. pylori* chromosomal DNA was isolated according to a basic DNA protocol outlined in Schleif R.F. and Wensink P.C., *Practical Methods in Molecular Biology*, p.98, Springer-Verlag, NY., 1981, with minor modifications. Briefly, cells were pelleted, resuspended in TE (10 mM Tris, 1 mM EDTA, pH 7.6) and GES lysis buffer (5.1 M guanidium thiocyanate, 0.1 M EDTA, pH 8.0, 0.5% N-laurylsarcosine) was added. Suspension was chilled and ammonium acetate ($NH_4Ac$) was added to final concentration of 2.0 M. DNA was extracted, first with chloroform, then with phenol-chloroform, and reextracted with chloroform. DNA was precipitated with isopropanol, washed twice with 70% EtOH, dried and resuspended in TE.

Following isolation whole genomic *H. pylori* DNA was nebulized (Bodenteich et al., *Automated DNA Sequencing and Analysis* (J.C. Venter, ed.), Academic Press, 1994) to a median size of 2000 bp. After nebulization, the DNA was concentrated and separated on a standard 1% agarose gel. Several fractions, corresponding to

approximate sizes 900-1300 bp, 1300-1700 bp, 1700-2200 bp, 2200-2700 bp, were excised from the gel and purified by the GeneClean procedure (Bio101, Inc.).

The purified DNA fragments were then blunt-ended using T4 DNA polymerase. The healed DNA was then ligated to unique BstXI-linker adapters in 100-1000 fold molar excess. These linkers are complimentary to the BstXI-cut pMPX vectors, while the overhang is not self-complimentary. Therefore, the linkers will not concatemerize nor will the cut-vector religate itself easily. The linker-adopted inserts were separated from the unincorporated linkers on a 1% agarose gel and purified using GeneClean. The linker-adopted inserts were then ligated to each of the 20 pMPX vectors to construct a series of "shotgun" subclone libraries. The vectors contain an out-of-frame lacZ gene at the cloning site which becomes in-frame in the event that an adapter-dimer is cloned, allowing these to be avoided by their blue-color.

All subsequent steps were based on the multiplex DNA sequencing protocols outlined in Church G.M. and Kieffer-Higgins S., *Science* 240:185-188, 1988. Only major modifications to the protocols are highlighted. Briefly, each of the 20 vectors was then transformed into DH5α competent cells (Gibco/BRL, DH5α transformation protocol). The libraries were assessed by plating onto antibiotic plates containing ampicillin, methicillin and IPTG/Xgal. The plates were incubated overnight at 37°C. Successful transformants were then used for plating of clones and pooling into the multiplex pools. The clones were picked and pooled into 40 ml growth medium cultures. The cultures were grown overnight at 37°C. DNA was purified using the Qiagen Midi-prep kits and Tip-100 columns (Qiagen, Inc.). In this manner, 100 μg of DNA was obtained per pool. Fifteen 96-well plates of DNA were generated to obtain a 5-10 fold sequence redundancy assuming 250-300 base average read-lengths.

These purified DNA samples were then sequenced using the multiplex DNA sequencing based on chemical degradation methods (Church G.M. and Kieffer-Higgins S., *Science* 240:185-188, 1988) or by Sequithrem (Epicenter Technologies) dideoxy sequencing protocols. The sequencing reactions were electrophoresed and transferred onto nylon membranes by direct transfer electrophoresis from 40 cm gels (Richterich P. and Church G.M., *Methods in Enzymology* 218:187-222, 1993) or by electroblotting (Church, *supra*). 24 samples were run per gel. 45 successful membranes were produced by chemical sequencing and 8 were produced by dideoxy sequencing. The DNA was covalently bound to the membranes by exposure to ultraviolet light, and hybridized with labeled oligonucleotides complimentary to tag sequences on the vectors (Church, *supra*). The membranes were washed to rinse off non-specifically bound probe, and exposed to X-ray film to visualize individual sequence ladders. After autoradiography, the hybridized probe was removed by incubation at 65° C, and the hybridization cycle

repeated with another tag sequence until the membrane had been probed 38 times for chemical sequencing membranes and 10 times for the dideoxy sequencing membranes. Thus, each gel produced a large number of films, each containing new sequencing information. Whenever a new blot was processed, it was initially probed for an internal standard sequence added to each of the pools.

Digital images of the films were generated using a laser-scanning densitometer (Molecular Dynamics, Sunnyvale, CA). The digitized images were processed on computer workstations (VaxStation 4000's) using the program REPLICA™ (Church et al., *Automated DNA Sequencing and Analysis* (J.C. Venter, ed.), Academic Press, 1994). Image processing included lane straightening, contrast adjustment to smooth out intensity differences, and resolution enhancement by iterative gaussian deconvolution. The sequences were then automatically picked in REPLICA™ and displayed for interactive proofreading before being stored in a project database. The proofreading was accomplished by a quick visual scan of the film image followed by mouse clicks on the bands of the displayed image to modify the base calls. Many of the sequence errors could be detected and corrected because multiple sequence reads covering the same portion of the genomic DNA provide adequate sequence redundancy for editing. Each sequence automatically received an identification number (corresponding to microtiter plate, probe information, and lane set number). This number serves as a permanent identifier of the sequence so it is always possible to identify the original of any particular sequence without recourse to a specialized database.

Routine assembly of *H. pylori* sequences was done using the program FALCON (Church, Church et al., *Automated DNA Sequenicng and Analysis* (J.C. Venter, ed.), Academic Press, 1994). This program has proven to be fast and reliable for most sequences. The assembled contigs were displayed using a modified version of GelAssemble, developed by the Genetics Computer Group (GCG) (Devereux et al., *Nucleic Acid Res.* 12:387-95, 1984) that interacts with REPLICA™. This provided for an integrated editor that allows multiple sequence gel images to be instantaneously called up from the REPLICA™ database and displayed to allow rapid scanning of contigs and proofreading of gel traces where discrepancies occurred between different sequence reads in the assembly.

II. Identification, cloning and expression of recombinant *H. pylori* DNA sequences

To facilitate the cloning, expression and purification of membrane and secreted proteins from *H. pylori* a powerful gene expression system, the pET System (Novagen), for cloning and expression of recombinant proteins in *E. coli*, was selected. Also, a DNA sequence encoding a peptide tag, the His-Tag, was fused to the 3' end of DNA

sequences of interest in order to facilitate purification of the recombinant protein products. The 3' end was selected for fusion in order to avoid alteration of any 5' terminal signal sequence. The exception to the above was ppiB, a gene cloned for use as a control in the expression studies. In this study, the sequence for *H. pylori* ppiB contains a DNA sequence encoding a His-Tag fused to the 5' end of the full length gene, because the protein product of this gene does not contain a signal sequence and is expressed as a cytosolic protein.

*PCR Amplification and cloning of DNA sequences containing ORF's for membrane and secreted proteins from the J99 Strain of Helicobacter pylori.*

Sequences chosen (from the list of the DNA sequences of the invention) for cloning from the J99 strain of *H. pylori* were prepared for amplification cloning by polymerase chain reaction (PCR). Synthetic oligonucleotide primers (Table 3) specific for the 5' and 3' ends of open reading frames (ORFs) were designed and purchased (GibcoBRL Life Technologies, Gaithersburg, MD, USA). All forward primers (specific for the 5' end of the sequence) were designed to include an NcoI cloning site at the extreme 5' terminus, except for HpSeq. 4821082 where NdeI was used. These primers were designed to permit initiation of protein translation at a methionine residue followed by a valine residue and the coding sequence for the remainder of the native *H. pylori* DNA sequence. An exception is *H. pylori* sequence 4821082 where the initiator methionine is immediately followed by the remainder of the native *H. pylori* DNA sequence. All reverse primers (specific for the 3' end of any *H. pylori* ORF) included a EcoRI site at the extreme 5' terminus to permit cloning of each *H. pylori* sequence into the reading frame of the pET-28b. The pET-28b vector provides sequence encoding an additional 20 carboxy-terminal amino acids (only 19 amino acids in HpSeq. 26380318 and HpSeq.14640637) including six histidine residues (at the extreme C-terminus), which comprise the His-Tag. An exception to the above, as noted earlier, is the vector construction for the ppiB gene. A synthetic oligonucleotide primer specific for the 5' end of ppiB gene encoded a BamHI site at its extreme 5' terminus and the primer for the 3' end of the ppiB gene encoded a XhoI site at its extreme 5' terminus.

TABLE 3

Oligonucleotide primers used for PCR amplification of *H. pylori* DNA sequences

| **Outer membrane Proteins** | **Forward primer 5' to 3'** | **Reverse Primer 5' to 3'** |
|---|---|---|
| Protein 16225006 | 5'-TATACCATGGTGGG CGCTAA-3' (SEQ ID NO:147) | 5'-ATGAATTCGAGTAAG GATTTTTG-3' (SEQ ID NO:148) |
| Protein 26054702 | 5'-TTAACCATGGTGAAA AGCGATA-3' (SEQ ID NO:149) | 5'-TAGAATTCGCATAAC GATCAATC-3' (SEQ ID NO:150) |
| Protein 7116626 | 5'-ATATCCATGGTGAGT TTGATGA-3' (SEQ ID NO:151) | 5'-ATGAATTCAATTTTT TATTTTGCCA-3' (SEQ ID NO:152) |
| Protein 29479681 | 5'-AATTCCATGGTGGGG GCTATG-3' (SEQ ID NO:153) | 5'-ATGAATTCTCGATAG CCAAAATC-3' (SEQ ID NO:154) |
| Protein 14640637 | 5'-AATTCCATGGTGCAT AACTTCCATT-3' (SEQ ID NO:155) | 5'-AAGAATTCTCTAGCA TCCAAATGGA-3' (SEQ ID NO:156) |
| **Periplasmic/ Secreted Proteins** | | |
| Protein 30100332 | 5'-ATTTCCATGGTCATG TCTCATATT-3' (SEQ ID NO:157) | 5'-ATGAATTCCATCTTT TATTCCAC-3' (SEQ ID NO:158) |
| Protein 4721061 | 5'-AACCATGGTGATTT TAAGCATTGAAAG-3' (SEQ ID NO:159) | 5'-AAGAATTCCACTCA AAATTTTTTAACAG-3' (SEQ ID NO:160) |
| **Other Surface Proteins** | | |
| Protein 4821082 | 5'-GATCATCCATATGTT ATCTTCTAAT-3' (SEQ ID NO:161) | 5'-TGAATTCAACCATTT TAACCCTG-3' (SEQ ID NO:162) |

| | | |
|---|---|---|
| Protein 978477 | 5'-TATACCATGGTGAA ATTTTTTCTTTTA-3' (SEQ ID NO:163) | 5'-AGAATTCAATTGCG TCTTGTAAAAG-3' (SEQ ID NO:164) |
| **Inner Membrane Protein** | | |
| Protein 26380318 | 5'-TATACCATGGTGAT GGACAAACTC-3' (SEQ ID NO:165) | 5'-ATGAATTCCCACTT GGGGCGATA-3' (SEQ ID NO:166) |
| **Cytoplasmic Protein** | | |
| ppi | 5'-TTATGGATCCAAAC CAATTAAAACT-3' (SEQ ID NO:167) | 5'-TATCTCGAGTTATA GAGAAGGGC-3' (SEQ ID NO:168) |

Genomic DNA prepared from the J99 strain of *H. pylori* (ATCC #55679; deposited by Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154) was used as the source of template DNA for PCR amplification reactions (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). To amplify a DNA sequence containing an *H. pylori* ORF, genomic DNA (50 nanograms) was introduced into a reaction vial containing 2 mM $MgCl_2$, 1 micromolar synthetic oligonucleotide primers (forward and reverse primers) complementary to and flanking a defined *H. pylori* ORF, 0.2 mM of each deoxynucleotide triphosphate; dATP, dGTP, dCTP, dTTP and 2.5 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems. Inc., Branchburg, NJ, USA) in a final volume of 100 microliters. The following thermal cycling conditions were used to obtain amplified DNA products for each ORF using a Perkin Elmer Cetus/ GeneAmp PCR System 9600 thermal cycler:

**Protein 26054702, Protein 7116626, Protein 29479681, Protein 30100332, and Protein 4821082;**

Denaturation at 94°C for 2 min,

2 cycles at 94°C for 15 sec, 30°C for 15 sec and 72°C for 1.5 min

23 cycles at 94°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min

Reactions were concluded at 72°C for 6 minutes.

**Protein 16225006;**

Denaturation at 94°C for 2 min.

25 cycles at 95°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min
Reaction was concluded at 72°C for 6 minutes.

**Protein 4721061;**
Denaturation at 94°C for 2 min,
2 cycles at 94°C for 15 sec, 36°C for 15 sec and 72°C for 1.5 min
23 cycles at 94°C for 15 sec, 60°C for 15 sec and 72°C for 1.5 min
Reactions were concluded at 72°C for 6 minutes.

**Protein 26380318;**
Denaturation at 94°C for 2 min,
2 cycles at 94°C for 15 sec, 38°C for 15 sec and 72°C for 1.5 min
23 cycles at 94°C for 15 sec, 62°C for 15 sec and 72°C for 1.5 min
Reactions were concluded at 72°C for 6 minutes.

**Protein 14640637;**
Denaturation at 94°C for 2 min,
2 cycles at 94°C for 15 sec, 33°C for 15 sec and 72°C for 1.5 min
30 cycles at 94°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min
Reactions were concluded at 72°C for 6 minutes.

**Conditions for amplification of *H. pylori* ppiB;**
Denaturation at 94°C for 2 min,
2 cycles at 94°C for 15 sec, 32°C for 15 sec and 72°C for 1.5 min
25 cycles at 94°C for 15 sec, 56°C for 15 sec and 72°C for 1.5 min
Reactions were concluded at 72°C for 6 minutes

Upon completion of thermal cycling reactions, each sample of amplified DNA was washed and purified using the Qiaquick Spin PCR purification kit (Qiagen, Gaithersburg, MD, USA). All amplified DNA samples were subjected to digestion with the restriction endonucleases, NcoI and EcoRI (New England BioLabs, Beverly, MA, USA), or in the case of HpSeq. 4821082 (SEQ ID NO: 1309), with NdeI and EcoRI (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). DNA samples were then subjected to electrophoresis on 1.0 % NuSeive (FMC BioProducts, Rockland, ME USA) agarose gels. DNA was visualized by exposure to ethidium bromide and long wave uv irradiation. DNA contained in slices

isolated from the agarose gel was purified using the Bio 101 GeneClean Kit protocol (Bio 101 Vista, CA, USA).

*Cloning of H. pylori DNA sequences into the pET-28b prokaryotic expression vector.*

The pET-28b vector was prepared for cloning by digestion with NcoI and EcoRI, or in the case of *H. pylori* protein 4821082 with NdeI and EcoRI (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). In the case of cloning ppiB, the pET-28a vector, which encodes a His-Tag that can be fused to the 5' end of an inserted gene, was used and the cloning site prepared for cloning with the ppiB gene by digestion with BamHI and XhoI restriction endonucleases.

Following digestion, DNA inserts were cloned (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994) into the previously digested pET-28b expression vector, except for the amplified insert for ppiB, which was cloned into the pET-28a expression vector. Products of the ligation reaction were then used to transform the BL21 strain of *E. coli* (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994) as described below.

*Transformation of competent bacteria with recombinant plasmids*

Competent bacteria, *E coli* strain BL21 or *E. coli* strain BL21(DE3), were transformed with recombinant pET expression plasmids carrying the cloned *H. pylori* sequences according to standard methods (Current Protocols in Molecular, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). Briefly, 1 microliter of ligation reaction was mixed with 50 microliters of electrocompetent cells and subjected to a high voltage pulse, after which, samples were incubated in 0.45 milliliters SOC medium (0.5% yeast extract, 2.0 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4 and 20, mM glucose) at 37°C with shaking for 1 hour. Samples were then spread on LB agar plates containing 25 microgram/ml kanamycin sulfate for growth overnight. Transformed colonies of BL21 were then picked and analyzed to evaluate cloned inserts as described below.

*Identification of recombinant pET expression plasmids carrying H. pylori sequences*

Individual BL21 clones transformed with recombinant pET-28b-H.pylori ORFs were analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers, specific for each *H. pylori* sequence, that were used in the original PCR amplification cloning reactions. Successful amplification verified the integration of the *H. pylori* sequences in the expression vector (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994).

*Isolation and Preparation of plasmid DNA from BL21 transformants*

Individual clones of recombinant pET-28b vectors carrying properly cloned *H. pylori* ORFs were picked and incubated in 5 mls of LB broth plus 25 microgram/ml kanamycin sulfate overnight. The following day plasmid DNA was isolated and purified using the Qiagen plasmid purification protocol (Qiagen Inc., Chatsworth, CA, USA).

*Expression of recombinant H. pylori sequences in E. coli*

The pET vector can be propagated in any *E. coli* K-12 strain e.g. HMS174, HB101, JM109, DH5, etc. for the purpose of cloning or plasmid preparation. Hosts for expression include *E. coli* strains containing a chromosomal copy of the gene for T7 RNA polymerase. These hosts are lysogens of bacteriophage DE3, a lambda derivative that carries the lacI gene, the lacUV5 promoter and the gene for T7 RNA polymerase. T7 RNA polymerase is induced by addition of isopropyl-B-D-thiogalactoside (IPTG), and the T7 RNA polymerase transcribes any target plasmid, such as pET-28b, carrying a T7 promoter and a gene of interest. Strains used include: BL21(DE3) (Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Meth. Enzymol. 185, 60-89).

To express recombinant *H. pylori* sequences, 50 nanograms of plasmid DNA isolated as described above was used to transform competent BL21(DE3) bacteria as described above (provided by Novagen as part of the pET expression system kit). The lacZ gene (beta-galactosidase) was expressed in the pET-System as described for the *H. pylori* recombinant constructions. Transformed cells were cultured in SOC medium for 1 hour, and the culture was then plated on LB plates containing 25 micrograms/ml kanamycin sulfate. The following day, bacterial colonies were pooled and grown in LB medium containing kanamycin sulfate (25 micrograms/ml) to an optical density at 600 nM of 0.5 to 1.0 O.D. units, at which point, 1 millimolar IPTG was added to the culture for 3 hours to induce gene expression of the *H. pylori* recombinant DNA constructions.

After induction of gene expression with IPTG, bacteria were pelleted by centrifugation in a Sorvall RC-3B centrifuge at 3500 x g for 15 minutes at 4°C. Pellets were resuspended in 50 milliliters of cold 10 mM Tris-HCl, pH 8.0, 0.1 M NaCl and 0.1 mM EDTA (STE buffer). Cells were then centrifuged at 2000 x g for 20 min at 4°C. Wet pellets were weighed and frozen at -80°C until ready for protein purification.

III. Purification of recombinant proteins from *E. coli*

*Analytical Methods*

The concentrations of purified protein preparations were quantified spectrophotometrically using absorbance coefficients calculated from amino acid

content (Perkins, S.J. 1986 Eur. J. Biochem. 157, 169-180). Protein concentrations were also measured by the method of Bradford, M.M. (1976) Anal. Biochem. 72, 248-254, and Lowry, O.H., Rosebrough, N., Farr, A.L. & Randall, R.J. (1951) J. Biol. Chem. 193, pages 265-275, using bovine serum albumin as a standard.

SDS-polyacrylamide gels (12% or 4.0 to 25 % acrylamide gradient gels) were purchased from BioRad (Hercules, CA, USA), and stained with Coomassie blue. Molecular weight markers included rabbit skeletal muscle myosin (200 kDa), *E. coli* (-galactosidase (116 kDa), rabbit muscle phosphorylase B (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa), egg white lysozyme (14.4 kDa) and bovine aprotinin (6.5 kDa).

*1. Purification of soluble proteins*

All steps were carried out at 4°C. Frozen cells were thawed, resuspended in 5 volumes of lysis buffer (20 mM Tris, pH 7.9, 0.5 M NaCl, 5 mM imidazole with 10% glycerol, 0.1 % 2-mercaptoethanol, 200 µg/ ml lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 10 ug/ml each of leupeptin, aprotinin, pepstatin, L-1-chloro-3-[4-tosylamido]-7-amino-2-heptanone (TLCK), L-1-chloro-3-[4-tosylamido]-4-phenyl-2-butanone (TPCK), and soybean trypsin inhibitor, and ruptured by several passages through a small volume microfluidizer (Model M-110S, Microfluidics International Corporation, Newton, MA). The resultant homogenate was made 0.1 % Brij 35, and centrifuged at 100,000 x g for 1 hour to yield a clear supernatant (crude extract).

Following filtration through a 0.8 µm Supor filter (Gelman Sciences, FRG) the crude extract was loaded directly onto a $Ni^{2+}$- nitrilotriacetate-agarose (NTA) with a 5 milliliter bed volume (Hochuli, E., Dbeli, H., and Schacheer, A. (1987) J. Chromatography 411, 177-184) pre-equilibrated in lysis buffer containing 10 % glycerol, 0.1 % Brij 35 and 1 mM PMSF. The column was washed with 250 ml (50 bed volumes) of lysis buffer containing 10 % glycerol, 0.1 % Brij 35, and was eluted with sequential steps of lysis buffer containing 10 % glycerol, 0.05 % Brij 35, 1 mM PMSF, and 20, 100, 200, and 500 mM imidazole in succession. Fractions were monitored by absorbance at $OD_{280}$ nm, and peak fractions were analyzed by SDS-PAGE. Fractions containing the recombinant protein eluted at 100 mM imidazole.

*Recombinant protein 14640637 and proteins, beta-galactosidase (lacZ) and peptidyl-prolyl cis-trans isomerase (ppiB)*

Fractions containing the recombinant proteins from the $Ni^{2+}$-NTA-agarose columns were pooled and then concentrated to approximately 5 ml by centrifugal

filtration (Centriprep-10, Amicon, MA), and loaded directly onto a 180-ml column (1.6 X 91 cm) of Sephacryl S-100 HR gel filtration medium equilibrated in Buffer A (10 mM Hepes, pH 7.5, 150 mM NaCl, 0.1 mM EGTA) and run in Buffer A at 18 ml/h. Fractions containing the recombinant protein were identified by absorbance at 280 nm and analyzed by SDS-PAGE. Fractions were pooled and concentrated by centrifugal filtration.

*Recombinant protein 7116626*

Fractions containing the recombinant protein from the $Ni^{2+}$ -NTA-agarose column were pooled and dialyzed overnight against 1 liter of dialysis buffer (10 mM MOPS, pH 6.5, 50 mM NaCl, 0.1 mM EGTA, 0.02% Brij 35 and 1 mM PMSF). In the morning, a fine white precipitate was removed by centrifugation and the resulting supernatant was loaded onto an 8 ml (8 x 75 mm) MonoS high performance liquid chromatography column (Pharmacia Biotechnology, Inc., Piscataway, NJ, USA) equilibrated in buffer B (10 mM MOPS, pH 6.5, 0.1 mM EGTA) containing 50 mM NaCl. The column was washed with 10 bed volumes of buffer B containing 50 mM NaCl, and developed with a 50-ml linear gradient of increasing NaCl (50 to 500 mM). Recombinant protein 7116626 eluted as a sharp peak at 300 mM NaCl.

*2. Purification of insoluble proteins from inclusion bodies*

The following steps were carried out at 4°C. Cell pellets were resuspended in lysis buffer with 10% glycerol 200 µg/ ml lysozyme, 5 mM EDTA, 1mM PMSF and 0.1 % -mercaptoethanol. After passage through the cell disrupter, the resulting homogenate was made 0.2 % deoxycholate, stirred 10 minutes, then centrifuged at 20,000 x g, for 30 min. The pellets were washed with lysis buffer containing 10 % glycerol, 10 mM EDTA, 1% Triton X-100, 1 mM PMSF and 0.1% -mercaptoethanol, followed by several washes with lysis buffer containing 1 M urea, 1 mM PMSF and 0.1 % 2-mercaptoethanol. The resulting white pellet was composed primarily of inclusion bodies, free of unbroken cells and membranous materials..

*Recombinant proteins 26054702, 16225006, 30100332, 4721061*

The following steps were carried out at room temperature. Purified inclusion bodies were dissolved in 20 ml 8.0 M urea in lysis buffer with 1 mM PMSF and 0.1 % 2-mercaptoethanol, and incubated at room temperature for 1 hour. Materials that did not dissolve were removed by centrifugation. The clear supernatant was filtered, then loaded onto a $Ni^{2+}$ -NTA agarose column pre-equilibrated in 8.0 M urea in Lysis Buffer. The column was washed with 250 ml (50 bed volumes) of lysis buffer

containing 8 M urea, 1.0 mM PMSF and 0.1 % 2-mercaptoethanol, and developed with sequential steps of lysis buffer containing 8M urea, 1 mM PMSF, 0.1 % 2-mercaptoethanol and 20, 100, 200, and 500 mM imidazole in succession. Fractions were monitored by absorbance at $OD_{280}$ nm, and peak fractions were analyzed by SDS-PAGE. Fractions containing the recombinant protein eluted at 100 mM imidazole.

*Recombinant proteins 29479681, 26380318*

The pellet containing the inclusion bodies was solubilized in buffer B containing 8 M urea, 1 mM PMSF and 0.1 % 2-mercaptoethanol, and incubated for 1 hour at room temperature. Insoluble materials were removed by centrifugation at 20,000 x g for 30 min, and the cleared supernatant was loaded onto a 15 ml ( 1.6 x 7.5 cm ) SP-Sepharose column pre-equilibrated in buffer B, 6 M urea, 1 mM PMSF, 0.1 % 2-mercaptoethanol. After washing the column with 10 bed volumes, the column was developed with a linear gradient from 0 to 500 mM NaCl.

*Dialysis and concentration of protein samples*

Urea was removed slowly from the protein samples by dialysis against Tris-buffered saline (TBS; 10 mM Tris pH 8.0, 150 mM NaCl) containing 0.5 % deoxycholate (DOC) with sequential reduction in urea concentration as follows; 6M, 4M, 3M, 2M, 1M, 0.5 M and finally TBS without any urea. Each dialysis step was conducted for a minimum of 4 hours at room temperature.

After dialysis, samples were concentrated by pressure filtration using Amicon stirred-cells. Protein concentrations were measured using the methods of Perkins (1986 Eur. J. Biochem. 157, 169-180), Bradford ((1976) Anal. Biochem. 72, 248-254) and Lowry ((1951) J. Biol. Chem. 193, pages 265-275).

The recombinant proteins purified by the methods described above are summarized in Table 4 below.

TABLE 4

| J99 Sequence Identifier | Homolog identified by Blast | Gene symbol of Homolog | Bacterial cell fraction used to purify recombinant proteins | Method of purification | Relative MW on SDS-PAGE gel | Final concentration of purified protein | Composition of buffer |
|---|---|---|---|---|---|---|---|
| | | | | | | | |
| **Outer Membrane Proteins** | | | | | | | |
| | | | | | | | |
| 16225006 | P28635 | **YEAC** | Inclusion bodies | His-Tag | 18 kDa | 5 mg/ml | B |
| | | | | | | | |
| 26054702 | P15929 | **flgH** | Inclusion bodies | His-Tag | 37 kDa | 1.18 mg/ml | B |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | ---- | as dry pellet |
| | | | | | | | |
| 7116626 | P26093 | e(P4) | Soluble fraction | His-Tag | 29 kDa | 0.8 mg/ml | A |
| | | | | | | 1.85 mg/ml | C |
| | | | | | | | |
| 29479681 | P13036 | fecA | Inclusions bodies | SP-Sepharose | 23 kDa | 2.36 mg/ml | B |
| | | | | | | 0.5 mg ml | B |
| | | | | | | ---- | as dry pellet |
| | | | | | | | |
| 14640637 | P16665 | TPF1 | Soluble fraction | His-Tag | 17 kDa | 2.4 mg/ml | A |
| | | | | gel filtration S100 HR | | | |
| | | | | | | | |

**Periplasmic/Secreted Protein**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 3010032 | P23847 | dppA | Inclusion bodies | His-Tag | 11 kDa | 2.88 mg/ml | B |
| | | | | | | | |
| 4721061 | P36175 | GCP | Inclusion bodies | His-Tag | 38 kDa | 2.8 mg/ml | B |
| | | | | | | | |

**Other Surface Proteins**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 4821082 | P08089 | M protein | Inclusion bodies | His-Tag | 20 kDa | 1.16 mg/ml | B |
| | | | | | | | |
| 978477 | L28919 | FBP54 | Inclusion bodies | SP-Sepharose | 44 kDa | 2.56 mg/ml | B |
| | | | | | | 0.3 mg/ml | B |

**Inner Membrane Proteins**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 26380318 | P15933 | fliG | Inclusion bodies | SP-Sepharose | 11 kDa | 22 mg/ml | B |
| | | | | | | | |
| | | | | | | | |

**Control Proteins with His-Tag**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | P00722 | lacZ | Soluble fraction | His-Tag | 116 kDa | 10 mg/ml | A |
| | | | | gel filtration S200 HR | | | |
| | | | | | | | |
| | | ppiB | Soluble fraction | His-Tag | 21 kDa | 4.4 mg/ml | A |
| | | | | gel filtration S100 HR | | | |
| Buffer compositions: | | | | | | | |
| A=10 mM Hepes pH 7.5, 150 mM NaCl, 0.1 mM EGTA | | | | | | | |
| B= 10 mM Tris pH 8.0, 150 mM NaCl, 0.5 % DOC | | | | | | | |
| C= 10 mM MOPS pH 6.5, 300 mM NaCl, 0.1 EGTA | | | | | | | |
| | | | | | | | |

5

IV. Analysis of *H. pylori* proteins as Vaccine candidates

To investigate the immunomodulatory effect of *H. pylori* proteins, a mouse/*H. pylori* model was used. This model mimics the human *H. pylori* infection in many respects. The focus is on the effect of oral immunization in *H. pylori* infected animals in order to test the concept of therapeutic oral immunotherapy.

*Animals*

Female SPF BALB/c mice were purchased from Bomholt Breeding center (Denmark). They were kept in ordinary makrolon cages with free supply of water and food. The animals were 4-6 weeks old at arrival.

*Infection*

After a minimum of one week of acclimatization, the animals were infected with a type 2 strain (VacA negative) of *H. pylori* (strain 244, originally isolated from an ulcer patient). In our hands, this strain has earlier proven to be a good colonizer of the mouse stomach. The bacteria were grown overnight in Brucella broth supplemented with 10 % fetal calf serum, at 37°C in a microaerophilic atmosphere (10% $CO_2$, 5%$O_2$). The animals were given an oral dose of omeprazole (400 µmol/kg) and 3-5 h after this an oral inoculation of *H. pylori* in broth (approximately $10^8$ cfu/animal). Positive take of the infection was checked in some animals 2-3 weeks after the inoculation.

*Antigens*

Recombinant *H. pylori* antigens were chosen based on their association with externally exposed *H. pylori* cell membrane. These antigens were selected from the following groups: (1.) Outer Membrane Proteins; (2.) Periplastic/Secreted proteins; (3.) Outer Surface proteins; and (4.) Inner Membrane proteins. All recombinant proteins were constructed with a hexa-HIS tag for purification reasons and the non-*Helicobacter pylori* control protein (b-galactosidase from *E. coli*; LacZ), was constructed in the same way.

All antigens were given in a soluble form, i.e. dissolved in either a HEPES buffer or in a buffer containing 0.5% Deoxycholate (DOC).

The antigens are listed in Table 5 below.

Table 5
*Helicobacter pylori* proteins

**Outer membrane Proteins**
Protein 7116626
Protein 4721061

Protein 16225006
Protein 29479681
Protein 14640637

**Periplasmic/Secreted Proteins**
Protein 30100332

**Other cell envelope proteins**
Protein 4821082

**Flagella-associated proteins**
Protein 26380318

**Control proteins**
b-galactosidase (LacZ)

*Immunizations*

Ten animals in each group were immunized 4 times over a 34 day period (day 1, 15, 25 and 35). Purified antigens in solution or suspension were given at a dose of 100 mg/mouse. As an adjuvant, the animals were also given 10 μg/mouse of Cholera toxin (CT) with each immunization. Omeprazole (400 mmol/kg) was given orally to the animals 3-5 h prior to immunization as a way of protecting the antigens from acid degradation. Infected control animals received HEPES buffer + CT or DOC buffer + CT. Animals were sacrificed 2-4 weeks after final immunization. A general outline of the study is shown in Table 6 below.

Table 6
Study outline, therapeutic immunization:

Mice were all infected with *H. pylori* strain Ah244 at day 30.

| Substance | Mouse strain n=10 | Dose/mouse | Dates for dosing |
|---|---|---|---|
| 1. Controls, PBS | Balb/c | 0.3 ml | 0, 14, 24, 34 |
| 2. Cholera toxin, 10 μg | Balb/c | 0.3 ml | 0, 14, 24, 34 |
| 3. Protein 16225006, 100 μg + CT 10 μg | Balb/c | 0.3 ml | 0, 14, 24, 34 |
| 4. Protein 26054702, 100 μg + CT 10 μg | Balb/c | 0.3 ml | 0, 14, 24, 34 |

| | | |
|---|---|---|
| 5. Protein 26380318, 100 µg + CT 10 µg Balb/c | 0.3 ml | 0, 14, 24, 34 |
| 6. Protein 29479681, 100 µg + CT 10 µg Balb/c | 0.3 ml | 0, 14, 24, 34 |
| 7. Protein 30100332, 100 µg + CT 10 µg Balb/c | 0.3 ml | 0, 14, 24, 34 |
| 8. Protein 4721061, 100 µg + CT 10 µg Balb/c | 0.3 ml | 0, 14, 24, 34 |
| 9. Protein 4821082, 100 µg + CT 10 µg Balb/c | 0.3 ml | 0, 14, 24, 34 |
| 10. Protein 7116626, 100 µg + CT 10 µg Balb/c | 0.3 ml | 0, 14, 24, 34 |
| 11.Protein 14640637, 100 µg + CT 10 µg Balb/c | 0.3 ml | 0, 14, 24, 34 |

*Analysis of infection*

Mucosal infection: The mice were sacrificed by $CO_2$ and cervical dislocation. The abdomen was opened and the stomach removed. After cutting the stomach along the greater curvature, it was rinsed in saline. The mucosa from the antrum and corpus of an area of $25mm^2$ was scraped separately with a surgical scalpel. The mucosa scraping was suspended in Brucella broth and plated onto Blood Skirrow selective plates. The plates were incubated under microaerophilic conditions for 3-5 days and the number of colonies was counted. The identity of *H. pylori* was ascertained by urease and catalase test and by direct microscopy or Gram staining.

The urease test was performed essentially as follows. The reagent, Urea Agar Base Concentrate, was purchased from DIFCO Laboratories, Detroit, MI (Catalog # 0284-61-3). Urea agar base concentrate was diluted 1:10 with water. 1 ml of if the diluted concentrate was mixed with 100-200 ml of actively growing *H. pylori* cells. Color change to magenta indicated that cells were urease positive.

The catalase test was performed essentially as follows. The reagent, N,N,N',N'-Tetramethyl-p-Phenylenediamine, was purchased from Sigma, St. Louis, MO (Catalog # T3134). A solution of the regent (1% w/v in water) was prepared. *H. pylori* cells were swabbed onto Whatman filter paper and overlaid with the 1% solution. Color change to dark blue indicated that the cells were catalase positive.

Serum antibodies: From all mice serum was prepared from blood drawn by heart puncture. Serum antibodies were identified by regular ELISA techniques, where the specific antigens of *Helicobacter pylori* were plated.

Mucosal antibodies: Gentle scrapings of a defined part of the corpus and of 4 cm of duodenum were performed in 50% of the mice in order to detect the presence of

antibodies in the mucous. The antibody titers were determined by regular ELISA technique as for serum antibodies.

Statistical analysis: Wilcoxon-Mann-Whitney sign rank test was used for determination of significant effects of the antigens on *Helicobacter pylori* colonization. $P<0.05$ was considered significant. Because the antrum is the major colonization site for *Helicobacter* most emphasis was put upon changes in the antral colonization.

*Results*

Antibodies in sera: All antigens tested given together with CT gave rise to a measurable specific titer in serum. The highest responses were seen with Protein 7116626, Protein 4721061, Protein 26380318, Protein 14640637 and Protein 4821082 (see Figure 1).

Antibodies in mucus: In the mucus scrapings, specific antibodies against all antigens tested were seen. By far the strongest response was seen with Protein 30100332, followed by Protein 14640637, and Protein 26380318 (see Figure 2).

Therapeutic immunization effects:

All control animals (BALB/c mice) were well colonized with *H. pylori* (strain AH244) in both antrum and corpus of the stomach. Of the antigens tested 3 proteins (Protein 4721061, Protein 4821082, and Protein 14640637) gave a good and significant reduction and/or eradication of the *H. pylori* infection. The degree of colonization of the antrum was lower following immunization with Protein 7116626 and Protein 26380318 compared to control. The effect of Proteins 16225006, 29479681, and 30100332 did not differ from control. The control protein lacZ, i.e. the non-*H. pylori* protein, had no eradication effect and in fact had higher *Helicobacter* colonization compared to the HEPES + CT control. All data are shown in Figures 3 and 4 for proteins dissolved in HEPES and DOC respectively. Data is shown as geometric mean values. n=8-10 Wilcoxon-Mann-Whitney sign rank test * = $p<0.05$; x/10 = number of mice showing eradication of *H. pylori* over the total number of mice examined.

The data presented indicate that all of the *H. pylori* associated proteins included in this study, when used as oral immunogens in conjunction with the oral adjuvant CT, resulted in stimulation of an immune response as measured by specific serum and mucosal antibodies. A majority of the proteins led to a reduction, and in some cases complete clearance of the colonization of *H. pylori* in this animal model. It should be noted that the reduction or clearance was due to heterologous protection rather than homologous protection (the polypeptides were based on the *H. pylori* J99 strain

sequence and used in the therapeutic immunization studies against a different (AH244) challenge strain, indicating the vaccine potential against a wide variety of *H. pylori* strains.

The highest colonization in the antrum was seen in animals treated with the non-*Helicobacter* protein LacZ, indicating that the effects seen with the *Helicobacter pylori* antigens were specific.

Taken together these data strongly support the use of these *H. pylori* proteins in a pharmaceutical formulation for the use in humans to treat and/or prevent *H. pylori* infections.

V. Sequence Variance Analysis of genes in *Helicobacter pylori* strains

Four genes were cloned and sequenced from several strains of *H. pylori* to compare the DNA and deduced amino acid sequences. This information was used to determine the sequence variation between the *H. pylori* strain, J99, and other *H. pylori* strains isolated from human patients.

*Preparation of Chromosomal DNA.*

Cultures of *H. pylori* strains (as listed in Table 9) were grown in BLBB (1% Tryptone, 1% Peptamin 0.1% Glucose, 0.2% Yeast Extract 0.5% Sodium Chloride, 5% Fetal Bovine Serum) to an $OD_{600}$ of 0.2. Cells were centrifuged in a Sorvall RC-3B at 3500 x g at 4°C for 15 minutes and the pellet resuspended in 0.95 mls of 10 mM Tris-HCl, 0.1 mM EDTA (TE). Lysozyme was added to a final concentration of 1mg/ml along with, SDS to 1% and RNAse A + T1 to 0.5mg/ml and 5 units/ml respectively, and incubated at 37°C for one hour. Proteinase K was then added to a final concentration of 0.4mg/ml and the sample was incubated at 55 C for more than one hour. NaCl was added to the sample to a concentration of 0.65 M, mixed carefully, and 0.15 ml of 10% CTAB in 0.7M NaCL (final is 1% CTAB/70mM NaCL) was added followed by incubation at 65°C for 20 minutes. At this point, the samples were extracted with chloroform:isoamyl alcohol, extracted with phenol, and extracted again with chloroform:isoamyl alcohol. DNA was precipitated with either EtOH (1.5 x volumes) or isopropanol (0.6 x volumes) at -70°C for 10minutes, washed in 70% EtOH and resuspended in TE.

*PCR Amplification and cloning.*

Genomic DNA prepared from twelve strains of *Helicobacter pylori* was used as the source of template DNA for PCR amplification reactions (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). To

amplify a DNA sequence containing an *H. pylori* ORF, genomic DNA (10 nanograms) was introduced into a reaction vial containing 2 mM $MgCl_2$, 1 micromolar synthetic oligonucleotide primers (forward and reverse primers, see Table 7) complementary to and flanking a defined *H. pylori* ORF, 0.2 mM of each deoxynucleotide triphosphate; dATP, dGTP, dCTP, dTTP and 0.5 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA) in a final volume of 20 microliters in duplicate reactions.

Table 7

Oligonucleotide primers used for PCR amplification of *H. pylori* DNA sequences.

| **Outer membrane Proteins** | **Forward primer 5' to 3'** | **Reverse Primer 5' to 3'** |
|---|---|---|
| **Protein 26054702** (for strains AH4, AH15, AH61, 5294, 5640, AH18, and AH244) | 5'-TTAACCATGGTGAAAAGCGATA-3' (SEQ ID NO:169) | 5'-TAGAATTCGCCTCTAAAACTTTAG-3' (SEQ ID NO:170) |
| **Protein 26054702** (for strains AH5, 5155, 7958, AH24,and J99) | 5'-TTAACCATGGTGAAAAGCGATA-3' (SEQ ID NO:171) | 5'-TAGAATTCGCATAACGATCAATC-3' (SEQ ID NO:172) |
| **Protein 7116626** | 5'-ATATCCATGGTGAGTTTGATGA-3' (SEQ ID NO:173) | 5'-ATGAATTCAATTTTTTATTTTGCCA-3' (SEQ ID NO:174) |
| **Protein 29479681** | 5'-AATTCCATGGCTATCCAAATCCG-3' (SEQ ID NO:175) | 5'-ATGAATTCGCCAAAATCGTAGTATT-3' (SEQ ID NO:176) |
| **Protein 346** | 5'-GATACCATGGAATTTATGAAAAAG-3' (SEQ ID NO:177) | 5'-TGAATTCGAAAAAGTGTAGTTATAC-3' (SEQ ID NO:178) |

The following thermal cycling conditions were used to obtain amplified DNA products for each ORF using a Perkin Elmer Cetus/ GeneAmp PCR System 9600 thermal cycler:

Protein 7116626 and Protein 346;
Denaturation at 94°C for 2 min,
2 cycles at 94°C for 15 sec, 30°C for 15 sec and 72°C for 1.5 min
23 cycles at 94°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min
Reactions were concluded at 72°C for 6 minutes.

Protein 26054702 for strains AH5, 5155, 7958, AH24,and J99;
Denaturation at 94°C for 2 min,
2 cycles at 94°C for 15 sec, 30°C for 15 sec and 72°C for 1.5 min
25 cycles at 94°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min
Reaction was concluded at 72°C for 6 minutes.

Protein 26054702 and Protein 294796813 for strains AH4, AH15, AH61, 5294, 5640, AH18, and Hp244 ;
Denaturation at 94°C for 2 min,
2 cycles at 94°C for 15 sec, 30°C for 20 sec and 72°C for 2 min
25 cycles at 94°C for 15 sec, 55°C for 20 sec and 72°C for 2 min
Reactions were concluded at 72°C for 8 minutes.

Upon completion of thermal cycling reactions, each pair of samples were combined and used directly for cloning into the pCR cloning vector as described below.

*Cloning of H. pylori DNA sequences into the pCR TA cloning vector.*

All amplified inserts were cloned into the pCR 2.1 vector by the method described in the Original TA cloning kit (Invitrogen, San Diego, CA). Products of the ligation reaction were then used to transform the TOP10F' (INVaF' in the case of *H. pylori* sequence 350) strain of *E. coli* as described below.

*Transformation of competent bacteria with recombinant plasmids*

Competent bacteria, *E coli* strain TOP10F' or *E. coli* strain INVaF' were transformed with recombinant pCR expression plasmids carrying the cloned *H. pylori* sequences according to standard methods (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). Briefly, 2 microliters of 0.5 micromolar BME was added to each vial of 50 microliters of competent cells. Subsequently, 2 microliters of ligation reaction was mixed with the competent cells and incubated on ice for 30 minutes. The cells and ligation mixture were then subjected to a

"heat shock" at 42°C for 30 seconds, and were subsequently placed on ice for an additional 2 minutes, after which, samples were incubated in 0.45 milliliters SOC medium (0.5% yeast extract, 2.0 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4 and 20, mM glucose) at 37°C with shaking for 1 hour. Samples were then spread on LB agar plates containing 25 microgram/ml kanamycin sulfate or 100 micrograms/ml ampicillan for growth overnight. Transformed colonies of TOP10F' or INVaF' were then picked and analyzed to evaluate cloned inserts as described below.

*Identification of recombinant PCR plasmids carrying H. pylori sequences*

Individual TOP10F' or INVaF' clones transformed with recombinant pCR-*H.pylori* ORFs were analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers, specific for each *H. pylori* sequence, that were used in the original PCR amplification cloning reactions. Successful amplification verified the integration of the *H. pylori* sequences in the cloning vector (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994).

Individual clones of recombinant pCR vectors carrying properly cloned *H. pylori* ORFs were picked for sequence analysis. Sequence analysis was performed on ABI Sequencers using standard protocols (Perkin Elmer) using vector-specific primers (as found in PCRII or pCR2.1, Invitrogen, San Diego, CA) and sequencing primers specific to the ORF as listed in Table 8 below.

Table 8
Oligonucleotide primers used for sequencing of *H. pylori* DNA sequences.

| Outer membrane Proteins | Forward primers 5' to 3' | Reverse Primers 5' to 3' |
|---|---|---|
| **Protein 26054702** | 5'-CCCTTCATTTTAGAAATCG-3' (SEQ ID NO:179)<br>5'-ATTTCAACCAATTCAATGCG-3' (SEQ ID NO:180)<br>5'-GCCCCTTTTGATTTGAAGCT-3' (SEQ ID NO:181)<br>5'-TCGCTCCAAGATACCAAGAAGT-3' (SEQ ID NO:182)<br>5'-CTTGAATTAGGGGCAAAGATCG-3' (SEQ ID NO:183)<br>5'-ATGCGTTTTTACCCAAAGAAGT-3' (SEQ ID NO:184)<br>5'-ATAACGCCACTTCCTTATTGGT-3' (SEQ ID NO:185) | 5'-CTTTGGGTAAAAACGCATC-3' (SEQ ID NO:186)<br>5'-CGATCTTTGATCCTAATTCA-3' (SEQ ID NO:187)<br>5'-ATCAAGTTGCCTATGCTGA-3' (SEQ ID NO:188) |
| | | |
| **Protein 7116626** | 5'-TTGAACACTTTTGATTATGCGG-3' (SEQ ID NO:189)<br>5'-GGATTATGCGATTGTTTTACAAG-3' (SEQ ID NO:190) | 5'-GTCTTTAGCAAAAATGGCGTC-3' (SEQ ID NO:191)<br>5'-AATGAGCGTAAGAGAGCCTTC-3' (SEQ ID NO:192) |
| **Protein 29479681** | 5'-CTTATGGGGGTATTGTCA-3' (SEQ ID NO:193)<br>5'-AGCATGTGGGTATCCAGC-3' (SEQ ID NO:194) | 5'-AGGTTGTTGCCTAAAGACT-3' (SEQ ID NO:195)<br>5'-CTGCCTCCACCTTTGATC-3' (SEQ ID NO:196) |

| | | |
|---|---|---|
| **Protein 346** | 5'-ACCAATATCAATTGGCACT-3' (SEQ ID NO:197)<br>5'-ACTTGGAAAAGCTCTGCA-3' (SEQ ID NO:198) | 5'-CTTGCTTGTCATATCTAGC-3' (SEQ ID NO:199)<br>5'-GTTGAAGTGTTGGTGCTA-3' (SEQ ID NO:200) |
| | 5'-CAAGCAAGTGGTTTGGTTTTAG-3' (SEQ ID NO:201)<br>5'-TGGAAAGAGCAAATCATTGAAG-3' (SEQ ID NO:202) | 5'-GCCCATAATCAAAAAGCCCAT-3' (SEQ ID NO:203)<br>5'-CTAAAACCAAACCACTTGCT<br>TGTC-3' (SEQ ID NO:204) |
| | | |
| **Vector Primers** | 5'-GTAAAACGACGGCCAG-3' (SEQ ID NO:205) | 5'-CAGGAAACAGCTATGAC-3' (SEQ ID NO:206) |

*Results*

To establish the PCR error rate in these experiments, five individual clones of Protein 26054702, prepared from five separate PCR reaction mixtures from *H. pylori* strain J99, were sequenced over a total length of 897 nucleotides for a cumulative total of 4485 bases of DNA sequence. DNA sequence for the five clones was compared to a DNA sequence obtained previously by a different method, i.e., random shotgun cloning and sequencing. The PCR error rate for the experiments described herein was determined to be 2 base changes out of 4485 bases, which is equivalent to an estimated error rate of less than or equal to 0.04%.

DNA sequence analysis was performed on four different open reading frames identified as genes and amplified by PCR methods from a dozen different strains of the bacterium *Helicobacter pylori*. The deduced amino acid sequences of three of the four open reading frames that were selected for this study showed statistically significant BLAST homology to defined proteins present in other bacterial species. Those ORFs included: Protein 26054702, homologous to the val A & B genes encoding an ABC transporter in F. novicida; Protein 7116626, homologous to lipoprotein e (P4) present in the outer membrane of H. influenzae; Protein 29479681, homologous to fecA, an outer membrane receptor in iron (III) dicitrate transport in *E. coli*. Protein 346 was identified as an unknown open reading frame, because it showed low homology with sequences in the public databases.

To assess the extent of conservation or variance in the ORFs across various strains of *H. pylori*, changes in DNA sequence and the deduced protein sequence were compared to the DNA and deduced protein sequences found in the J99 strain of *H.*

*pylori* (*see* Table 9 below). Results are presented as percent identity to the J99 strain of *H. pylori* sequenced by random shotgun cloning. To control for any variations in the J99 sequence each of the four open reading frames were cloned and sequenced again from the J99 bacterial strain and that sequence information was compared to the sequence information that had been collected from inserts cloned by random shotgun sequencing of the J99 strain. The data demonstrate that there is variation in the DNA sequence ranging from as little as 0.12 % difference (Protein 346, J99 strain) to approximately 7% change (Protein 26054702, strain AH5). The deduced protein sequences show either no variation ( Protein 346, strains AH18 and AH24) or up to as much as 7.66% amino acid changes (Protein 26054702, Strain AH5).

**Table 9**

Multiple Strain DNA Sequence analysis of H. pylori Vaccine Candidates

| J99 Protein #: | 26054702 | | 2054702 | | 7116626 | | 29479681 | | 29479681 | | 346 | | 346 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Length of Region Sequenced: | 248 a.a. | | 746 nt. | | 232 a.a. | | 96 nt. | | 182 a.a. | | 548 nt. | | 273 a.a. | 819 nt. |
| Strain Tested | AA identity | Nuc. identity | AA identity | Nuc. identity | AA identity | Nuc. identity | AA identity | Nuc. identity | AA identity | Nuc. identity | AA identity | Nuc. identity | AA identity | Nuc. identity |
| J99 | 100.00% | 100.00% | 100.00% | 100.00% | 100.00% | 100.00% | 99.63% | 99.88% | | | | | | |
| AH244 | 95.16% | 95.04% | n.d. | n.d. | 99.09% | 96.71% | 98.90% | 96.45% | | | | | | |
| AH4 | 95.97% | 95.98% | 97.84% | 95.83% | n.d. | n.d. | 97.80% | 95.73% | | | | | | |
| AH5 | 92.34% | 93.03% | 98.28% | 96.12% | 98.91% | 96.90% | 98.53% | 95.73% | | | | | | |
| AH15 | 95.16% | 94.91% | 97.41% | 95.98% | 99.82% | 97.99% | 99.63% | 96.09% | | | | | | |
| AH61 | n.d. | n.d. | 97.84% | 95.98% | 99.27% | 97.44% | n.d. | n.d. | | | | | | |
| 5155 | n.d. | n.d. | n.d. | n.d. | 99.45% | 97.08% | 98.53% | 95.60% | | | | | | |
| 5294 | 94.35% | 94.37% | 98.28% | 95.40% | 99.64% | 97.26% | 97.07% | 95.48% | | | | | | |
| 7958 | 94.35% | 94.10% | 97.84% | 95.40% | n.d. | n.d. | 99.63% | 96.46% | | | | | | |
| 5640 | 95.16% | 94.37% | 97.41% | 95.69% | 99.09% | 97.63% | 98.53% | 95.48% | | | | | | |
| AH18 | n.d. | n.d. | 98.71% | 95.69% | 99.64% | 97.44% | 100.00% | 95.97% | | | | | | |
| AH24 | 94.75% | 95.04% | 97.84% | 95.40% | 99.27% | 96.71% | 100.00% | 96.46% | | | | | | |

n.d.= not done.

VI. Experimental Knock-Out Protocol for the Determination of Essential *H. pylori* Genes as Potential Therapeutic Targets

Therapeutic targets are chosen from genes whose protein products appear to play key roles in essential cell pathways such as cell envelope synthesis, DNA synthesis, transcription, translation, regulation and colonization/virulence.

The protocol for the deletion of portions of *H. pylori* genes/ORFs and the insertional mutagenesis of a kanamycin-resistance cassette in order to identify genes which are essential to the cell is modified from previously published methods (Labigne-Roussel et al., 1988, J. Bacteriology 170, pp. 1704-1708; Cover et al.,1994, J. Biological Chemistry 269, pp. 10566-10573; Reyrat et al., 1995, Proc. Natl. Acad. Sci. 92, pp 8768-8772). The result is a gene "knock-out."

*Identification and Cloning of H. pylori Gene Sequences*

The sequences of the genes or ORFs (open reading frames) selected as knock-out targets are identified from the *H. pylori* genomic sequence and used to design primers to specifically amplify the genes/ORFs. All synthetic oligonucleotide primers are designed with the aid of the OLIGO program (National Biosciences, Inc., Plymouth, MN 55447, USA), and can be purchased from Gibco/BRL Life Technologies (Gaithersburg, MD, USA). If the ORF is smaller than 800 to 1000 base pairs, flanking primers are chosen outside of the open reading frame.

Genomic DNA prepared from the *Helicobacter pylori* HpJ99 strain (ATCC 55679; deposited by Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154) is used as the source of template DNA for amplification of the ORFs by PCR (polymerase chain reaction) (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). For the preparation of genomic DNA from *H. pylori*, see Example I. PCR amplification is carried out by introducing 10 nanograms of genomic HpJ99 DNA into a reaction vial containing 10 mM Tris pH 8.3, 50 mM KCl, 2 mM $MgCl_2$, 2 microMolar synthetic oligonucleotide primers (forward=F1 and reverse=R1), 0.2 mM of each deoxynucleotide triphosphate (dATP,dGTP, dCTP, dTTP), and 1.25 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA) in a final volume of 40 microliters. The PCR is carried out with Perkin Elmer Cetus/GeneAmp PCR System 9600 thermal cyclers.

Upon completion of thermal cycling reactions, each sample of amplified DNA is visualized on a 2% TAE agarose gel stained with Ethidium Bromide (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994) to

determine that a single product of the expected size had resulted from the reaction. Amplified DNA is then washed and purified using the Qiaquick Spin PCR purification kit (Qiagen, Gaithersburg, MD, USA).

PCR products are cloned into the pT7Blue T-Vector (catalog#69820-1, Novagen, Inc., Madison, WI, USA) using the TA cloning strategy (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). The ligation of the PCR product into the vector is accomplished by mixing a 6 fold molar excess of the PCR product, 10 ng of pT7Blue-T vector (Novagen), 1 microliter of T4 DNA Ligase Buffer (New England Biolabs, Beverly, MA, USA), and 200 units of T4 DNA Ligase (New England Biolabs) into a final reaction volume of 10 microliters. Ligation is allowed to proceed for 16 hours at 16°C.

Ligation products are electroporated (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994) into electroporation-competent XL-1 Blue or DH5-a *E.coli* cells (Clontech Lab., Inc. Palo Alto, CA, USA). Briefly, 1 microliter of ligation reaction is mixed with 40 microliters of electrocompetent cells and subjected to a high voltage pulse (25 microFarads, 2.5 kV, 200 ohms) after which the samples are incubated in 0.45 ml SOC medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose) at 37°C with shaking for 1 hour. Samples are then spread onto LB (10 g/l bacto tryptone, 5 g/l bacto yeast extract, 10 g/l sodium chloride) plates containing 100 microgram/ml of Ampicillin, 0.3% X-gal, and 100 microgram/ml IPTG. These plates are incubated overnight at 37°C. Ampicillin-resistant colonies with white color are selected, grown in 5 ml of liquid LB containing 100 microgram/ml of Ampicillin, and plasmid DNA is isolated using the Qiagen miniprep protocol (Qiagen, Gaithersburg, MD, USA).

To verify that the correct *H.pylori* DNA inserts had been cloned, these pT7Blue plasmid DNAs are used as templates for PCR amplification of the cloned inserts, using the same forward and reverse primers used for the initial amplification of the J99 *H.pylori* sequence. Recognition of the primers and a PCR product of the correct size as visualized on a 2% TAE, ethidium bromide stained agarose gel are confirmation that the correct inserts had been cloned. Two to six such verified clones are obtained for each knock-out target, and frozen at -70°C for storage. To minimize errors due to PCR, plasmid DNA from these verified clones are pooled, and used in subsequent cloning steps.

The sequences of the genes/ORFs are again used to design a second pair of primers which flank the region of *H. pylori* DNA to be either interrupted or deleted (up to 250 basepairs) within the ORFs but are oriented away from each other. The pool of

circular plasmid DNAs of the previously isolated clones are used as templates for this round of PCR. Since the orientation of amplification of this pair of deletion primers is away from each other, the portion of the ORF between the primers is not included in the resultant PCR product. The PCR product is a linear piece of DNA with *H. pylori* DNA at each end and the pT7Blue vector backbone between them which, in essence, resultes in the deletion of a portion of the ORFs. The PCR product is visualized on a 1% TAE, ethidium bromide stained agarose gel to confirm that only a single product of the correct size has been amplified.

A Kanamycin-resistance cassette (Labigne-Roussel et al., 1988 J. Bacteriology 170, 1704-1708) is ligated to this PCR product by the TA cloning method used previously (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). The Kanamycin cassette containing a *Campylobacter* kanamycin resistance gene is obtained by carrying out an EcoRI digestion of the recombinant plasmid pCTB8:*kan* (Cover et al.,1994, J. Biological Chemistry 269, pp. 10566-10573). The proper fragment (1.4 kb) is isolated on a 1% TAE gel, and isolated using the QIAquick gel extraction kit (Qiagen, Gaithersburg, MD, USA). The fragment is end repaired using the Klenow fill-in protocol, which involved mixing 4ug of the DNA fragment, 1 microliter of dATP,dGTP, dCTP, dTTP at 0.5 mM, 2 microliter of Klenow Buffer (New England Biolabs) and 5 units of Klenow DNA Polymerase I Large (Klenow) Fragment (New England Biolabs) into a 20 microliter reaction, incubating at 30°C for 15 min, and inactivating the enzyme by heating to 75°C for 10 minutes. This blunt-ended Kanamycin cassette is then purified through a Qiaquick column (Qiagen, Gaithersburg, MD, USA) to eliminate nucleotides. The "T" overhang is then generated by mixing 5 micrograms of the blunt-ended kanamycin cassette, 10 mM Tris pH 8.3, 50 mM KCl, 2 mM $MgCl_2$, 5 units of DNA Polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA), 20 microliters of 5 mM dTTP, in a 100 microliter reaction and incubating the reaction for 2 hours at 37°C. The "Kan-T" cassette is purified using a QIAquick column (Qiagen, Gaithersburg, MD, USA). The PCR product of the deletion primers (F2 and R2) is ligated to the Kan-T cassette by mixing 10 to 25 ng of deletion primer PCR product, 50 - 75 ng Kan-T cassette DNA, 1 microliter 10x T4 DNA Ligase reaction mixture, 0.5 microliter T4 DNA Ligase (New England Biolabs, Beverly, MA, USA) in a 10 microliter reaction and incubating for 16 hours at 16°C.

The ligation products are transformed into XL-1 Blue or DH5-a *E.coli* cells by electroporation as described previously. After recovery in SOC, cells are plated onto LB plates containing 100 microgram/ml Ampicillin and grown overnight at 37°C. These plates are then replica plated onto plates containing 25 microgram/ml Kanamycin and

allowed to grow overnight. Resultant colonies have both the Ampicillin resistance gene present in the pT7Blue vector, and the newly introduced Kanamycin resistance gene. Colonies are picked into LB containing 25 microgram/ml Kanamycin and plasmid DNA is isolated from the cultured cells using the Qiagen miniprep protocol (Qiagen, Gaithersburg, MD, USA).

Several tests by PCR amplification are conducted on these plasmids to verify that the Kanamycin is inserted in the *H. pylori* gene/ORF, and to determine the orientation of the insertion of the Kanamycin-resistance gene relative to the *H. pylori* gene/ORF. To verify that the Kanamycin cassette is inserted into the *H. pylori* sequence, the plasmid DNAs are used as templates for PCR amplification with the set of primers originally used to clone the *H. pylori* gene/ORFs. The correct PCR product is the size of the deleted gene/ORF but increased in size by the addition of a 1.4 kilobase Kanamycin cassette. To avoid potential polar effects of the kanamycin resistance cassette on *H. pylori* gene expression, the orientation of the Kanamycin resistance gene with respect to the knock-out gene/ORF is determined and both orientations are eventually used in *H. pylori* transformations (see below). To determine the orientation of insertion of the kanamycin resistance gene, primers are designed from the ends of the kanamycin resistance gene ("Kan-1" 5'-ATCTTACCTATCACCTCAAAT-3' (SEQ ID NO:207)), and "Kan-2" 5'-AGACAGCAACATCTTTGTGAA-3' (SEQ ID NO:208)). By using each of the cloning primers in conjunction with each of the Kan primers (4 combinations of primers), the orientation of the Kanamycin cassette relative to the *H.pylori* sequence is determined. Positive clones are classified as either in the "A" orientation (the same direction of transcription is present for both the *H. pylori* gene and the Kanamycin resistance gene), or in the "B" orientation (the direction of transcription for the *H.pylori* gene is opposite to that of the Kanamycin resistance gene). Clones which share the same orientation (A or B) are pooled for subsequent experiments and independently transformed into *H. pylori*.

*Transformation of Plasmid DNA into H. pylori cells*

Two strains of *H. pylori* are used for transformation: ATCC 55679, the clinical isolate which provided the DNA from which the *H. pylori* sequence database is obtained, and AH244, an isolate which had been passaged in, and has the ability to colonize the mouse stomach. Cells for transformation are grown at 37°C, 10% $CO_2$, 100% humidity, either on Sheep-Blood agar plates or in Brucella Broth liquid. Cells are grown to exponential phase, and examined microscopically to determine that the cells are "healthy" (actively moving cells) and not contaminated. If grown on plates, cells are harvested by scraping cells from the plate with a sterile loop, suspended in 1 ml of

Brucella Broth, spun down (1 minute, top speed in eppendorf microfuge) and resuspended in 200 microliters Brucella Broth. If grown in Brucella Broth liquid, cells are centrifuged (15 minutes at 3000 rpm in a Beckman TJ6 centrifuge) and the cell pellet resuspended in 200 microliters of Brucella broth. An aliquot of cells is taken to determine the optical density at 600 nm, in order to calculate the concentration of cells. An aliquot (1 to 5 $OD_{600}$ units/25 microliter) of the resuspended cells is placed onto a prewarmed Sheep-Blood agar plate, and the plate is further incubated at 37°C, 6% $CO_2$, 100% humidity for 4 hours. After this incubation, 10 microliters of plasmid DNA (100 micrograms per microliter) is spotted onto these cells. A positive control (plasmid DNA with the ribonuclease H gene disrupted by kanamycin resistance gene) and a negative control (no plasmid DNA) are done in parallel. The plates are returned to 37°C, 6% $CO_2$ for an additional 4 hours of incubation. Cells are then spread onto that plate using a swab wetted in Brucella broth, and grown for 20 hours at 37°C, 6% $CO_2$. Cells are then transferred to a Sheep-Blood agar plate containing 25 micrograms/ml Kanamycin, and allowed to grow for 3 to 5 days at 37°C, 6% $CO_2$, 100% humidity. If colonies appear, they are picked and regrown as patches on a fresh Sheep-Blood agar plate containing 25 micrograms/ml Kanamycin.

Three sets of PCR tests are done to verify that the colonies of transformants have arisen from homologous recombination at the proper chromosomal location. The template for PCR (DNA from the colony) is obtained by a rapid boiling DNA preparation method as follows. An aliquot of the colony (stab of the colony with a toothpick) is introduced into 100 microliters of 1% Triton X-100, 20 mM Tris, pH 8.5, and boiled for 6 minutes. An equal volume of phenol : chloroform (1:1) is added and vortexed. The mixture is microfuged for 5 minutes and the supernatant is used as DNA template for PCR with combinations of the following primers to verify homologous recombination at the proper chromosomal location.

TEST 1. PCR with cloning primers originally used to amplify the gene/ORF. A positive result of homologous recombination at the correct chromosomal location should show a single PCR product whose size is expected to be the size of the deleted gene/ORF but increased in size by the addition of a 1.4 kilobase Kanamycin cassette. A PCR product of just the size of the gene/ORF is proof that the gene had not been knocked out and that the transformant is not the result of homologous recombination at the correct chromosome location.

TEST 2. PCR with F3 (primer designed from sequences upstream of the gene/ORF and not present on the plasmid), and either primer Kan-1 or Kan-2 (primers designed from the ends of the kanamycin resistance gene), depending on whether the plasmid DNA used was of "A" or "B" orientation. Homologous recombination at the

correct chromosomal location will result in a single PCR product of the expected size (i.e., from the location of F3 to the insertion site of kanamycin resistance gene). No PCR product or PCR product(s) of incorrect size(s) will prove that the plasmid had not integrated at the correct site and that the gene had not been knocked out.

TEST 3. PCR with R3 (primer designed from sequences downstream of the gene/ORF and not present on the plasmid) and either primer Kan-1 or Kan-2, depending on whether the plasmid DNA used was of "A" or "B" orientation. Homologous recombination at the correct chromosomal location will result in a single PCR product of the expected size (i.e., from the insertion site of kanamycin resistance gene to the downstream location of R3). Again, no PCR product or PCR product(s) of incorrect size(s) will prove that the plasmid had not integrated at the correct site and that the gene had not been knocked out.

Transformants showing positive results for all three tests above indicate that the gene is not essential for survival *in vitro*.

A negative result in any of the three above tests for each transformant indicates that the gene had not been disrupted, and that the gene is essential for survival *in vitro*.

In the event that no colonies result from two independent transformations while the positive control with the disrupted ribonuclease H plasmid DNA produces transformants, the plasmid DNA is further analyzed by PCR on DNA from transformant populations prior to plating for colony formation. This will verify that the plasmid can enter the cells and undergo homologous recombination at the correct site. Briefly, plasmid DNA is incubated according to the transformation protocol described above. DNA is extracted from the *H. pylori* cells immediately after incubation with the plasmid DNAs and the DNA is used as template for the above TEST 2 and TEST 3. Positive results in TEST 2 and TEST 3 would verify that the plasmid DNA could enter the cells and undergo homologous recombination at the correct chromosomal location. If TEST 2 and TEST 3 are positive, then failure to obtain viable transformants indicates that the gene is essential, and cells suffering a disruption in that gene are incapable of colony formation

VII. High-throughput drug screen assay

*Cloning, expression and protein purification*

Cloning, transformation, expression and purification of the *H. pylori* target gene and its protein product,e.g., an *H. pylori* enzyme, to be used in a high-throughput drug screen assay, is carried out essentially as described in Examples II and III above. Development and application of a screening assay for a particular *H. pylori* gene product, peptidyl-propyl *cis-trans* isomerase, is described below as a specific example.

*Enzymatic Assay*

The assay is essentially as described by Fisher (Fischer, G., et.al. (1984) *Biomed. Biochim. Acta* 43:1101-1111). The assay measures the *cis-trans* isomerization of the Ala-Pro bond in the test peptide N-succinyl-Ala-Ala-Pro-Phe-p-nitroanilide (Sigma # S-7388, lot # 84H5805). The assay is coupled with α-chymotrypsin, where the ability of the protease to cleave the test peptide occurs only when the Ala-Pro bond is in *trans*. The conversion of the test peptide to the trans isomer in the assay is followed at 390 nm on a Beckman Model DU-650 spectophotometer. The data are collected every second with an average scanning of time of 0.5 second. Assays are carried out in 35 mM Hepes, pH 8.0, in a final volume of 400 ul, with 10 μM α-chymotrypsin (type 1-5 from bovine Pancreas, Sigma # C-7762, lot 23H7020) and 10 nM PPIase. To initiate the reaction, 10 μl of the substrate ( 2 mM N-Succinyl-Ala-Ala-Pro-Phe-p-nitroanilide in DMSO) is added to 390 μl of reaction mixture at room temperature.

*Enzymatic assay in crude bacterial extract.*

A 50 ml culture of *Helicobacter pylori* (strain J99) in Brucella broth is harvested at mid-log phase ($OD_{600\ nm} \sim 1$) and resuspended in lysis buffer with the following protease inhibitors: 1 mM PMSF, and 10 μg/ml of each of aprotinin, leupeptin, pepstatine, TLCK, TPCK, and soybean trypsin inhibitor. The suspension is subjected to 3 cycles of freeze-thaw (15 minutes at -70 ° C, then 30 minutes at room temperature), followed by sonication (three 20 second bursts). The lysate is centrifuged (12,000 g x 30 minutes) and the supernatant is assayed for enzymatic activity as described above.

Many *H. pylori* enzymes can be expressed at high levels and in an active form in *E. coli*. Such high yields of purified proteins provide for the design of various high throughput drug screening assays.

## EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments and methods described herein. Such equivalents are intended to be encompassed by the scope of the following claims.

SEQUENCE LISTING

1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Astra Aktiebolag
(B) STREET: S-151 85
(C) CITY: Sodertalje
(D) STATE:
(E) COUNTRY: Sweden
(F) POSTAL CODE (ZIP)

(ii) TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO HELICOBACTER PYLORI AND VACCINE COMPOSITIONS THEREOF

(iii) NUMBER OF SEQUENCES: 208

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE:
(B) COMPUTER:
(C) OPERATING SYSTEM:
(D) SOFTWARE:

(v) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER
(B) FILING DATE:

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER:US 08/739,150
(B) FILING DATE: 28-OCT-1996

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/759,739
(B) FILING DATE: 06-DEC-1996

(viii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/891,928
(B) FILING DATE: 14-JULY-1997

(ix) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: LAHIVE & COCKFIELD
(B) STREET: 28 State Street
(C) CITY: Boston
(D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02109-1875

(x) ATTORNEY/AGENT INFORMATION:
(A) NAME: Mandragouras, Amy E.
(B) REGISTRATION NUMBER: 36,207
(C) REFERENCE/DOCKET NUMBER: GTN-001CP10PC

(xi) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617)227-7400
(B) TELEFAX: (617)742-4214

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 561 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...561

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGATTAAAA GAATTGCTTG TATTTTAAGC TTGAGCGCGA GTTTAGCGTT AGCTGGCGAA 60
GTGAATGGGT TTTTCATGGG TGCGGGTTAT CAACAAGGTC GTTATGGCCC TTATAACAGC 120
AATTACTCTG ATTGGCGTCA TGGCAATGAC CTTTATGGTT TGAATTTCAA ATTAGGTTTT 180
GTAGGCTTTG CCAATAAATG GTTTGGGGCT AGGGTGTATG GCTTTTTAGA TTGGTTTAAC 240
ACTTCAGGGA CTGAACACAC CAAAACCAAT TTGCTCACCT ATGGCGGCGG TGGCGATTTG 300
ATTGTCAATC TCATTCCTTT GGATAAATTC GCTCTAGGTC TCATTGGTGG CGTTCAATTA 360
GCCGGAAACA CTTGGATGTT CCCTTATGAT GTCAATCAAA CCAGATTCCA GTTCTTATGG 420
AATTTAGGCG GAAGAATGCG TGTTGGGGAT CGCAGTGCGT TTGAAGCGGG CGTGAAATTC 480
CCTATGGTTA ATCAGGGTAG CAAAGATGTA GGGCTTATCC GCTACTATTC TTGGTATGTG 540
GATTATGTCT TCACTTTCTA G 561

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 351 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...351

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
TTGATGCGCA TTATCATAAG GTTACTTTCA TTTAAAATGA ACGCTTTTTT AAAACTCGCG    60
CTCGCTTCTT TGATGGGGGG GCTTTGGTAT GCTTTCAATG GCGAAGGCTC TGAGATTGTC   120
GCTATAGGGA TTTTTGTGTT GATCTTGTTT GTTTTTTTTA TCCGCCCTGT GAGTTTCCAA   180
GACCCAGAAA AACGAGAAGA ATACATAGAA CGGCTTAAAA AAAACCATGA GAGGAAAATG   240
ATCTTACAAG ACAAGCAAAA AGAAGAGCAA ATGCGCCTCT ATCAAGCCAA AAAAGAGCGA   300
GAGAGCAGGC AAAAACAAGA CCTTAAAGAA CAAATGAAAA AATACTCATA A            351
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1038 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1038

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
ATGGTTAAAC ACTATCTTTT CATGGCGGTT TCGCAGGTCT TTTTCTCCTT CTTTTTAGTG    60
CTGTTTTTTA TCTCTTCCAT TGTGTTATTA ATCAGTATTG CAAGCGTAAC GCTCGTGATT   120
AAAGTGAGCT TTTTGGATCT GGTGCAACTC TTTTTGTATT CCTTGCCAGG AACCATTTTT   180
TTTATTTTGC CGATCACTTT TTTTGCGGCT TGCGCTTTGG GGCTTTCAAG GCTTAGCTAT   240
GACCATGAAT TGTTAGTGTT TTTCTCTTTA GGGGTTTCGC CTAAAAAAAT GACTAAAGCG   300
TTTGTGCCTT TAAGTTTGTT AGTGAGCGCG ATTTTATTAG CGTTTTCGCT CATCTTAATC   360
CCCACTTCTA AGAGCGCTTA TTACGGGTTT TTGCGTCAAA AAAAAGACAA GATTGACATT   420
AACATCAGAG CGGGTGAATT CGGGCAAAAA TTAGGCGATT GGCTCGTGTA TGTGGATAAG   480
ACTGAAAACA ATTCCTATGA TAATTTGGTG CTTTTTTCTA ATAAAAGTCT CTCTCAAGAA   540
AGCTTTATTT TGGCTCAAAA AGGCAATATC AACAATCAAA ACGGCGTGTT TGAATTGAAT   600
TTGTATAACG GGCATGCGTA TTTCACTCAA GGCGATAAAA TGCGTAAGGT TGATTTTGAA   660
GAATTGCATT TGCGCAACAA GCTCAAGTCT TTCAATTCTA ATGATGCGGC TTATTTGCAA   720
GGCACGGATT ATTTGGGTTA TTGGAAAAAA GCCTTTGGTA AAAACGCTAA TAAAAATCAA   780
AAACGCCGTT TTTCTCAAGC GATCTTAGTT TCCTTGTTCC CTTTAGCGAG CGTGTTTTTA   840
ATCCCCTTAT TTGGCATCGC CAACCCGCGA TTCAAAACGA ATTGGAGTTA TTTCTATGTC   900
CTTGGAGCGG TTGGGGTTTA TTTTTTAATG GTGCATGTGA TTTCTACGGA TTTGTTTTTG   960
ATGACCTTTT TCTTCCCCTT TATTTGGGCG TTTATTTCTT ATTTATTGTT TAGAAAATTC  1020
ATTTTAAAGC GTTATTAA                                                1038
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 831 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...831

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
ATGAAGAAAA AAGCAAAAGT CTTTTGGTGT TGTTTTAAAA TGATTCGTTG GTTGTATTTG    60
GCGGTCTTTT TTTTGTTGAG CGTATCAGAC GCTAAAGAAA TCGCTATGCA ACGATTTGAC   120
AAACAAAACC ATAAGATTTT TGAAATCCTT GCGGATAAAG TGAGCGCCAA AGACAATGTG   180
ATAACCGCCT CAGGGAATGC GATCCTATTG AATTATGACG TGTATATTCT AGCGGATAAG   240
GTGCGTTATG ACACCAAGAC TAAAGAAGCG TTATTAGAAG GCAATATTAA GGTTTATAGG   300
GGCGAGGGCT TGCTCGTTAA AACCGATTAT GTGAAATTGA GTTTGAACGA AAAATATGAG   360
ATCATTTTCC CCTTTTATGT CCAAGACAGC GTGAGCGGGA TTTGGGTGAG CGCGGATATT   420
GCTAGCGGGA AGGATCAAAA ATATAAGATT AAAAACATGA GCGCTTCAGG GTGCAGCATT   480
GACAACCCCA TTTGGCATGT CAATGCGACT TCAGGCTCAT TTAACATGCA AAAATCGCAT   540
TTGTCAATGT GGAATCCTAA GATTTATGTC GGCGATATTC CTGTATTGTA TTTGCCCTAT   600
ATTTTCATGT CCACGAGCAA TAAAAGAACT ACCGGGTTTT TATACCCTGA GTTTGGCACT   660
TCCAACTTAG ACGGCTTTAT TTATTTGCAA CCCTTTTATT TAGCCCCCAA AAACTCATGG   720
GATATGACCT TTACCCCACA AATCCGTTAC AAAAGGGGTT TTGGCTTGAA TTTTGAAGCG   780
CGCTACATCA ACTCTAAGAC GCAGGTTTTT ATTCAATGCG CGCTATTTTA G            831
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 675 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...675

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATGATTAGAT TAAAAGGTTT GAATAAAACT TTAAAAACAA GCTTATTAGC TGGGGTTTTA 60
CTAGGTGCTA CTGCTCCCTT AATGGCAAAG CCTTTATTAA GCGATGAAGA CTTATTGAAA 120
CGAGTAAAAC TACACAATAT CAAAGAAGAT ACGCTGACTA GCTGTAATGC TAAGGTGGAC 180
GGCTCTCAAT ACTTGAATAG TGGTTGGAAT TTATCTAAAG AATTTCCGCA AGAATATAGA 240
GAAAAGATTT TTGAATGCGT AGAAGAAGAA AAACATAAAC AAGCCCTTAA TTTAATCAAT 300
AAAGAAGACA CTAAAGATAA AGAAGAACTT GCAAAAAAAA TCAAAGAAAT TAAAGAAAAA 360
GCTAAAGTTT TAAGGCAAAA ATTTATGGCT TTTGAAATGA AAGAACACTC TAAAGAATTC 420
CCAAATAAAA AGCAACTTCA AACCATGCTT GAGAACGCTT TTGATAATGG AGCTGAAAGT 480
TTTATTGATG ATTGGCACGA ACGCTTTGGG GGTATAAGTA GAGAGAATAC TTATAAAGCA 540
CTTGGCATTA AAGAATATAG TGATGAAGGA AAGATATTGC CTTTGGCGAA AGAAGTTATA 600
TTAGACAATA TAAAAAAGAT TTTGAAGAAA GCACTTATGA TACTAGACAA CCCTTATCTG 660
CTATGGCTAG TATGA 675

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1290 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1290

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATGCCATACG CCTTAAGAAA AAGATTTTTC AAACGCCTTT TATTGTTTTT TTTAATTGTT 60
TGTATGATAA ATTTGCATGC CAAAAGCTAT CTGTTTTCTC CTTTGCCCCC AGCGCACCAG 120
CAAATCATTA AGACAGAGCC TTGCTCTTTG GAGTGCTTGA AAGACTTGAT GCTGCAAAAT 180
CAAATCTTTT CTTTTGTATC CCAATACGAT GATAACAACC AAGATGAGAG CCTTAAAACT 240
TATTACAAGG ACATCTTAAA CAAACTCAAC CCCGTATTCA TCGCTTCTCA AACTCCAGCT 300
AAAGAAAGCT ATGAGCCTAA GATTGAATTA GCGATTTTAC TGCCTAAAAA GGTGGTGGGC 360
CGTTATGCGA TTTTAGTGAT GAACACCCTT TTAGCGTATT TGAACACCAG AAACAACGAT 420
TTCAATATCC AAGTCTTTGA CAGCGATGAA GAAAGCCCTG AAAAATTAGA AGAAACCTAT 480
AAAGAAATTG AAAAAGAAAA ATTCCCTTTT ATCATCGCTT TATTGACTAA AGAGGGCGTG 540
GAAAATTTGC TCCAAAATAC GACTATCAAT ACCCCTACTT ATGTGCCTAC GGTGAATAAA 600
ACGCAATTAG AAAATCATAC CGAGCTTTCT TTAAGCGAGC GCTTGTATTT TGGGGGGATT 660
GATTATAAAG AGCAATTAGG CATGCTCGCA ACTTTCATTA GCCCTAATTC GCCCGTGATT 720
GAATACGATG ATGATGGCCT GATAGGTGAA CGCTTGAGGC AAATCACGGA GTCTTTAAAC 780

GTTGAAGTCA AACACCAAGA AAACATTTCT TACAAACAAG CGACCAGTTT TTCTAAAAAT 840
TTTAGAAAAC ATGATGCGTT TTTTAAAAAT TCTACCTTAA TTTTGAACAC CCCTACCACT 900
AAAAGCGGTC TGATCCTTTC TCAAATAGGG CTTTTAGAGT ATAAGCCTCT TAAAATCCTT 960
TCCACACAAA TCAATTTCAA CCCCTCTTTA CTCTTGCTCA CCCAGCCTAA AGACAGGAAA 1020
AATTTATTCA TTGTCAATGC CTTGCAAAAC AGCGATGAAA CGCTGATAGA ATACGCTTCC 1080
TTATTAGAGA GCGATTTAAG GCATGATTGG GTGAATTATT CCAGCGCGAT AGGGCTAGAG 1140
ATGTTTTTAA ACACGCTAGA TCCGCATTTT AAAAAGTCTT TTCAAGAGAG TTTGGAAGAC 1200
AATCAAGTCC GTTACCACAA TCAAATTTAT CAGGCTTTAG GGTATTCTTT TGAGCCGATA 1260
AAAAACGAAA GCGAAACAAA AAAAGAATAA 1290

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1368 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1368

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTGTTAAAAT TTCAAAAATT ACCCTTATTG TTTGTTTCCA TTCTTTATAA TCAAAGCCCT 60
TTATTGGCTT TTGATTATAA GTTTAGTGGG GTAGCGGAAT CTGTTTCTAA AGTGGGGTTT 120
AACCATTCCA AACTCAATTC CAAAGAAGGG ATTTTCCCTA CAGCCACCTT TGTAACCGCC 180
ACGATCAAGC TTCAAGTGGA TTCCAATCTG CTCCCTAAAA ACATTGAAAA ACACAGCTTA 240
AAAATAGGCG TTGGGGGGAT TTTAGGAGCG CTCGCTTACG ATTCCACCAA AACGCTCATA 300
GACCAAGCCA CGCATCAAAT CTATGGCTCA GAACTTTTTT ACCTCATAGG GCGTTGGTGG 360
GGGTTTTTAG GCAACGCTCC TTGGAAAGAC TCCCTCATAG AATCTGACGC TCACACCCGT 420
AATTATGTGC TGTATAATTC CTATCTGTTT TATTCTTATG GCGATAAATT CCACCTAAAA 480
TTAGGGCGTT ATCTCTCTAA CATGGATTTT ATGAGTTCCT ACACACAGGG TTTTGAACTG 540
GATTATAAAA TCAATTCTAA AATAGCGTTA AAATGGTTTA GCTCTTTTGG GAGGGCGTTG 600
GCTTTTGGGC AATGGATACG GGATTGGTAT GCCCCTATTG TAACTGAAGA TGGCAGAAAA 660
GAAGTTTATG ATGGCATCCA TGCCGCGCAA CTCTATTTTT CTAGCAAGCA TGTTCAAGTC 720
ATGCCTTTTG CTTATTTTTC GCCTAAGATT TACGGAGCGC CCGGTGTTAA AATCCATATT 780
GATAGCAACC CGAAATTCAA AGGCTTAGGG TTAAGGGCTC AAACCACTAT TAATGTGATT 840
TTCCCTGTTT ATGCTAAAGA TTTATACGAT GTGTATTGGC GTAACTCTAA GATTGGCGAG 900
TGGGGCGCAT CGCTTTTGAT CCACCAACGC TTTGACTACA ACGAATTTAA CTTTGGCTTT 960
GGTTATTACC AAAATTTTGG CAACGCTAAC GCAAGGATTG GCTGGTATGG TAACCCCATC 1020
CCTTTTAATT ATAGAAATAA CAGCGTTTAT GGTGGGGTCT TCAGTAACGC TATTACCGCA 1080
GACGCCGTTT CTGGGTATGT CTTTGGTGGG GGGGTGTATA GAGGGTTTTT ATGGGGTATT 1140
TTAGGCAGAT ACACTTATGC CACTAGAGCG AGCGAAAGAT CCATCAACTT GAACTTGGGC 1200
TATAAATGGG GTTCTTTTGC TAGAGTTGAT GTGAATTTAG AATACTATGT GGTCAGCATG 1260
CACAACGGCT ATAGATTAGA CTATCTCACC GGCCCTTTCA ACAAAGCCTT TAAGGCTGAC 1320

GCACAAGATA GGAGTAACCT TATGGTTAGC ATGAAATTCT TTTTTTAA 1368

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 849 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...849

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATGGGGTGTT CGTTTATCTT TAAAAAAGTT AGGGTTTATT CTAAAATGTT GGTTGCTTTG 60
GGGCTTTCAA GCGTGTTGAT CGGTTGCGCG ATGAATCCAA GCGCTGAGAC AAAAAAACCA 120
AATGACGCCA AAAACCAACA ACCAGTTCAA ACTCATGAAA GAATGACAAC AAGTTCTGAA 180
CATGTTACGC CACTAGATTT TAATTACCCG GTGCATATTG TTCAAGCCCC ACAAAACCAT 240
CATGTTGTAG GTATTTTAAT GCCACGCATT CAAGTGAGCG ATAATCTAAA ACCCTATATT 300
GATAAGTTTC AAGACGCTTT AATTAATCAA ATCCAAACTA TTTTTGAAAA AAGAGGCTAT 360
CAAGTGTTGC GTTTTCAAGA TGAAAAAGCT TTGAATGTGC AAGATAAGAA AAAGATTTTT 420
TCCGTTTTGG ATTTGAAAGG GTGGGTAGGA ATCTTAGAAG ATTTGAAAAT GAATTTAAAA 480
GATCCCAATA GTCCCAATTT AGACACGCTA GTGGATCAAA GCTCAGGCTC TGTATGGTTT 540
AATTTTTATG AACCAGAAAG CAATCGTGTC GTCCATGATT TTGCTGTAGA AGTAGGAACT 600
TTTCAGGCAA TAACATACAC ATACACCTCT ACTAATAACG CTTCAGGAGG GTTTAATTCT 660
TCAAAAAGCG TTATCCATGA AAATTTGGAT AAGAATAGAG AAGACGCGAT ACACAAGATT 720
TTAAACAGAA TGTATGCGGT TGTCATGAAA AAAGCTGTAA CAGAACTTAC AAAAGAAAAT 780
ATCGCCAAAT ACAGAGACGC TATTGATAGA ATGAAAGGCT TTAAAAGTTC TATGCCTCAA 840
AAAAAGTAG 849

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 843 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...843

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATGAAACTGA GAGCAAGTGT TTTAATCGGT GTGGCAATTC TGTGCTTAAT TTTAAGTGCG 60
TGCAGTAACT ATGCGAAAAA AGTGGTGAAA CAAAAGAACC ATGTTTATAC GCCTGTGTAT 120
AATGAACTGA TAGAGAAGTA TAGTGAGATC CCCTTAAATG ACAAACTCAA AGACACACCA 180
TTCATGGTGC AAGTGAAGTT GCCAAATTAC AAGGACTATT TGTTGGATAA TAAACAAGTT 240
GTACTAACTT TCAAACTTGT TCACCATTCT AAAAAGATTA CGCTCATAGG CGATGCCAAT 300
AAGATCCTCC AATACAAGAA TTACTTCCAA GCTAACGGGG CAAGATCTGA CATTGATTTT 360
TACTTGCAAC CCACTTTGAA TCAAAAGGGT GTGGTGATGA TAGCGAGTAA CTACAATGAT 420
AATCCCAACA ACAAAGAAAA ACCACAGACC TTTGATGTGT TGCAAGGAAG TCAGCCAATG 480
CTAGGAGCTA ACACAAAAAA CTTGCATGGC TATGATGTGA GTGGAGCAAA CAACAAGCAA 540
GTGATCAATG AAGTGGCAAG AGAAAAAGCT CAGCTAGAAA AAATCAATCA GTATTACAAG 600
ACTCTCTTGC AAGACAAGGA ACAAGAATAT ACCACTAGGA AAAATAACCA ACGAGAAATT 660
TTAGAAACAT TGAGTAATCG TGCAGGTTAT CAAATGAGGC AGAATGTGAT TAGTTCTGAG 720
ATTTTTAAGA ATGGCAACTT GAACATGCAA GCCAAAGAAG AAGAAGTTAG GGAGAAGCTA 780
CAAGAAGAAA GAGAGAATGA ATACTTGCGC AATCAAATCA GAAGTTTGCT CAGTGGTAAG 840
TGA 843

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1179 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1179

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATGAGAAAAC TATTCATCCC ACTTTTATTA TTCAGCGCTT TAGAAGCGAA CGAGAAAAAC 60
GGCTTTTTCA TAGAAGCCGG CTTTGAAACT GGGCTATTAG AAGGCACACA AACGCAAGAA 120
AAAAGACACA CCACCACAAA AAACACTTAC GCAACTTACA ATTATTTACC CACAGACACG 180
ATTTTAAAAA GAGCGGCTAA TTTATTCACC AATGCCGAAG CGATTTCAAA ATTAAAATTC 240
TCATCTTTAT CCCCTGTTAG AGTGTTGTAT ATGTATAATG GTCAATTAAC TATAGAAAAC 300
TTCTTGCCTT ATAATTTAAA TAATGTTAAG CTTAGTTTTA CAGACGCTCA AGGCAATGTG 360
ATCGATCTAG GCGTGATAGA GACTATCCCC AAACACTCTA AGATTGTTTT GCCCGGAGAG 420

GCATTTGATA GTCTAAAAAT TGACCCCTAT ACTTTATTTC TTCCAAAAAT TGAAGCCACT 480
AGCACTTCTA TTTCTGACGC TAACACGCAG AGGGTGTTTG AAACGCTCAA TAAGATTAAG 540
ACAAATTTGG TCGTAAATTA TAGGAATGAA AACAAATTTA AAGATCACGA AAATCATTGG 600
GAAGCCTTTA CCCCACAAAC CGCAGAAGAA TTCACTAATT TAATGTTGAA CATGATCGCT 660
GTTTTAGACT CCCAATCTTG GGGCGATGCG ATCTTAAACG CTCCTTTTGA GTTCACTAAC 720
AGCCCAACAG ATTGCGATAA TGATCCTTCA AAATGCGTAA ATCCTGGGAC AAACGGGCTT 780
GTCAATTCTA AAGTCGATCA AAAATATGTG TTAAACAAAC AAGACATTGT CAATAAATTT 840
AAAAACAAAG CGGATCTTGA TGTAATTGTT TTAAAGGATT CAGGGGTTGT AGGGCTTGGG 900
AGTGATATTA CCCCTAGCAA CAATGATGAT GGCAAGCATT ATGGCCAGTT AGGGGTAGTA 960
GCTTCTGCTT TAGATCCTAA AAAACTCTTT GGCGATAACC TTAAGACTAT CAATTTAGAG 1020
GATTTAAGAA CCATCTTGCA TGAATTCAGC CACACTAAAG GCTATGGGCA TAACGGGAAT 1080
ATGACCTATC AAAGAGTGCC GGTAACGAAA GATGGTCAAG TGGAAAAGGA TAGTAATGGC 1140
AAGCCAAAAG ATTCTGATGG CCTCCCCTAT AATGTGTGT 1179

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 813 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...813

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATGAAAAAGT TTGTAGCTTT AGGGCTTCTA TCCGCGGTTT TAAGCTCTTC GTTGTTAGCC 60
GAAGGTGATG GTGTTTATAT AGGGACTAAT TATCAGCTTG GACAAGCCCG TTTGAATAGC 120
AATATTTATA ATACAGGGGA TTGCACAGGG AGTGTTGTAG GTTGCCCCCC AGGTCTTACC 180
GCTAATAAGC ATAATCCAGG AGGCACCAAT ATCAATTGGC ACTCCAAATA CGCTAATGGG 240
GCTTTGAATG GTTTTGGGTT GAATGTGGGT TATAAGAAAT TCTTCCAATT CAAGTCGCTA 300
GATATGACAA GCAAGTGGTT TGGTTTTAGA GTGTATGGGC TTTTTGATTA CGGGCATGCC 360
GATTTAGGTA AACAAGTTTA TGCACCTAAT AAAATCCAGT TGGATATGGT CTCTTGGGGT 420
GTGGGGAGCG ATTTGTTAGC TGATATTATT GATAAAGACA ACGCTTCTTT TGGTATTTTT 480
GGTGGGGTCG CTATCGGCGG TAACACTTGG AAAAGCTCTG CAGCAAACTA TTGGAAAGAG 540
CAAATCATTG AAGCCAAAGG TCCTGATGTT TGTACCCCTA CTTATTGTAA CCCTAATGCC 600
CCTTATAGCA CCAACACTTC AACCGTCGCT TTTCAAGTGT GGTTGAATTT TGGGGTGAGA 660
GCCAATATCT ACAAGCATAA TGGCGTGGAA TTTGGCGTGA GAGTGCCGCT ACTCATCAAT 720
AAATTTTTGA GCGCGGGTCC TAACGCTACT AACCTTTATT ACCATTTGAA ACGGGATTAT 780
TCGCTTTATT TGGGGTATAA CTACACTTTT TAA 813

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 423 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...423

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ATGCATCCTA TAATGTTTGC CTATATCGCT AACGCGCTCG CTCAAGCTAG AAAGATCAAC 60
GGAACACTTT GCATGGCGTT TCAAAAAATA TCTCAAGTCA AAGAATTAGG CATTGATAAA 120
GCAAAGAGTT TGATAGGCAA CCTTTCTCAA GTGATTATCT ACCCCACAAA AGATACTGAT 180
GAATTAATAG AATGTGGCGT CCCATTAAGC GATAGTGAAA TCAATTTCTT ACACAACACG 240
GACATGAGAG CCAGACAAGT GCTAGTAAAA AATATCGTTA CAAACGCTTC AGCTTTTATT 300
GAAATTGATT TAAAAAAGAT TTGCAAGAAC TACTTTATAT TCTTGATAGC AATGCTGGTA 360
ATAGAAAAAT CCTCAATGAT CTTAAAAAAG CAAACCAAGA AACTTATAAG GAAGAGTATT 420
TAA 423

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 771 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...771

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATGTTGGGGA GCGTCAAAAA AGCGGTTTTT AGGGTTTTGT GTTTGGGGGC GTTGTGTTTA 60
TGCGGGGGGT TAATGGCAGA GCAAGATCCT AAAGAGCTTA TATTTTCAGG TATAACTATT 120
TACACGGATA AAAATTTCAC TAGAGCTAAG AAATATTTTG AAAAAGCTTG CAAATCAAAC 180

GATGCTGATG GCTGTGCAAT CTTAAGAGAG GTTTATTCTA GTGGTAAAGC CATAGCGAGA 240
GAAAACGCAA GAGAGAGCAT TGAAAAAGCT CTTGAACACA CCGCTACTGC TAAAGTTTGT 300
AAATTAAACG ATGCTGAAAA ATGCAAGGAC TTAGCAGAGT TTTATTTTAA TGTAAACGAT 360
CTTAAAAATG CTTTAGAATA TTACTCTAAA TCTTGTAAGT TAAATAATGT TGAAGGGTGT 420
ATGCTGTCAG CAACTTTTTA TAACGATATG ATAAAGGGTT TGAAAAAAGA TAAAAAAGAT 480
CTAGAATATT ATTCTAAAGC TTGCGAGTTA AATAACGGTG GAGGGTGTTC TAAATTAGGA 540
GGGGATTATT TTTTTGGTGA AGGCGTAACA AAAGATTTCA AAAAAGCTTT TGAATATTCT 600
GCCAAAGCTT GTGAGTTGAA CGATGCTAAA GGGTGTTACG CTCTAGCAGC GTTTTATAAT 660
GAGGGTAAAG GCGTGGCAAA GGATGAAAAG CAAACGACAG AAAACCTTGA AAAGAGTTGC 720
AAGCTAGGAT TAAAAGAAGC ATGCGATATT CTCAAAGAAC AAAAACAATA A 771

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 729 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...729

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATGAAAAAAT TTTTTTCTCA ATCTTTGTTA GCTCTTATTA TCTCTATGAA TGCGGTATCT 60
GGCATGGATG GTAATGGCGT TTTTTTAGGG GCGGGTTATT TGCAAGGACA GGCGCAAATG 120
CATGCGGATA TTAATTCTCA AAAACAAGCC ACCAACGCTA CGATCAAAGG CTTTGACGCG 180
CTCTTGGGGT ATCAATTTTT CTTTGAAAAA CACTTTGGCT TACGCCTTTA TGGGTTTTTT 240
GACTACGCTC ATGCCAATTC TATTAAGCTT AAAAACCCTA ACTATAATAG CGAAGCGGCG 300
CAAGTGGCTA GTCAAATTCT TGGGAAACAA GAAATCAATC GTTTAACAAA CATTGCCGAT 360
CCCAGAACTT TTGAGCCGAA CATGCTCACT TATGGGGGGG CTATGGACGT GATGGTTAAT 420
GTCATCAATA ACGGCATCAT GAGTTTGGGG GCTTTTGGCG GGATACAATT GGCCGGCAAT 480
TCATGGCTTA TGGCGACACC GAGCTTTGAG GGCATTTTAG TGGAACAAGC CCTTGTGAGC 540
AAGAAAGCCA CTTCTTTCCA ATTTTTATTC AATGTGGGGG CTCGCTTAAG GATCTTAAAA 600
CATTCTAGCA TTGAAGCGGG CGTGAAATTC CCCATGCTAA AGAAAAACCC CTACATCACT 660
GCAAAAAATT TGGATATAGG GTTTAGGCGC GTGTATTCGT GGTATGTGAA TTACGTGTTC 720
ACTTTCTAG 729

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 804 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...804

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATGAACTACC CTAATCTACC TAACAGCGCT TTAGAGATAA GCGAACAGCC AGAAGTGAAA 60
GAAATCACTA ACGAGCTTTT AAAGCAATTA CAAAACGCTT TAAGGAGCAA CGCGCATTTT 120
AGCGAGCAAG TGGAATTAAG CCTTAAATGC ATCGTTAGGA TTTTAGAAGT GCTTTTGAGT 180
TTGGATTTTT TTAAGAATGC GAATGAGATT GATAGCAGTT TAAGAAATTC CATTGAGTGG 240
CTGACTAACG CCGGCGAGAG CTTGAAATTA AAAATGAAAG AATACGAGCG CTTTTTTAGC 300
GAGTTTAATA CGAGCATGCA TGCCAACGAG CAGGAAGTAA CCAATACCTT AAACGCTAAC 360
GCCGAGAACA TTAAAAGCGA AATTAAAAAG CTAGAAAATC AATTGATAGA AACCACGACA 420
AGACTTTTAA CGAGCTATCA AATCTTTTTA AACCAAGCCA GAGATAACGC TAACAACCAA 480
ATCACAAAAA ACAAAACCCA AAGCCTTGAA GCGATTACAC AAGCTAAAAA CAACGCTAAT 540
AATGAAATAA GCAACAATCA AACGCAAGCG ATAACTAATA TCACCGAAGC GAAAACGAAC 600
GCTAATAATG AAATAAGCAA CAATCAAACG CAAGCGATAA CTAACATTAA CGAAGCCAAA 660
GAAAGCGCTA CAACGCAAAT AAACGCCAAT AAGCAAGAAG CAATAAATAA CATCACGCAA 720
GAAAAAACCC AAGCCACAAG CGAGATCACC GAAGCGAAAA AGACCGATCA TTATCAAAAC 780
ATTGATTTTT TTGAGTTTGA ATAA 804

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1632 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1632

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTGATAGAGA CCATCCCCAA ACACTCTAAG ATTGTTTTAC CCGGGGAGGC GTTTGATAGT 60

TTAAAAGAGG CGTTTGATAA AATTGACCCC TATACTTTCT TTTTTCCAAA ATTTGAAGCC 120
ACTAGCACTT CTATTTCTGA TACTAACACG CAGAGGGTGT TTGAAACGCT CAATAACATT 180
AAAACAAATC TTATAATGAA ATATAGTAAT GAAAATCCAA ACAATTTCAA CACTTGTCCT 240
TACAATAATA ATGGTAATAC AAAAAATGAT TGTTGGCAAA ATTTCACCCC ACAAACCGCA 300
GAAGAATTCA CCAATTTAAT GTTGAACATG ATCGCTGTCT TAGACTCCCA ATCTTGGGGC 360
GATGCGATCT TAAACGCTCC TTTTGAATTC ACTAACAGCT CAACAGATTG CGATAGCGAT 420
CCTTCAAAAT GCGTAAATCC CGGAGTAAAT GGGCGTGTTG ATACTAAAGT CGATCAACAA 480
TATATACTCA ACAAACAAGG TATTATTAAT AATTTTAGAA AAAAAATAGA AATTGATGCG 540
GTTGTTTTAA AAAATTCAGG GGTTGTAGGG TTAGCCAATG GATATGGCAA TGATGGTGAA 600
TATGGCACAT TAGGGGTAGA AGCCTATGCT TTAGATCCTA AAAAACTCTT TGGCAACGAC 660
CTTAAGACTA TCAATTTAGA AGATTTAAGA ACCATCTTGC ATGAATTCAG CCACACTAAA 720
GGCTATGGGC ATAACGGGAA TATGACCTAT CAAAGAGTGC CGGTAACGAA AGATGGTCAA 780
GTGGAAAAGG ATAGTAATGG CAAGCCAAAA GATTCTGATG GCCTCCCCTA TAATGTGTGT 840
TCGCTTTATG GGGGATCCAA TCAGCCCGCT TTCCCTAGCA ACTACCCTAA TTCCATCTAT 900
CACAATTGTG CGGATGTCCC GGCTGGCTTT TTAGGGGTAA CAGCAGCGGT TTGGCAGCAG 960
CTCATCAATC AAAACGCCTT GCCGATCAAC TACGCTAACT TGGGGAGTCA AACAAACTAC 1020
AACCTAAACG CTAGTTTAAA CACGCAAGAT TTAGCCAATT CCATGCTCAG CACCATCCAA 1080
AAAACCTTTG TAACTTCTAG CGTTACCAAC CACCATTTTT CAAACGCATC GCAAAGTTTT 1140
AGAAGCCCTA TTTTAGGGGT TAACGCTAAA ATAGGCTATC AAAACTACTT TAATGATTTC 1200
ATAGGGTTGG CTTATTATGG CATCATCAAA TACAATTACG CTAAAGCTGT TAATCAAAAA 1260
GTCCAGCAAT TGAGCTATGG TGGGGGGATA GATTTGTTAT TGGATTTCAT CACCACTTAC 1320
TCCAATAAAA ATAGCCCTAC AGGCATTCAA ACCAAAAGGA ATTTTTCTTC ATCTTTTGGT 1380
ATCTTTGGGG GGTTAAGGGG CTTGTATAAC AGCTATTATG TGTTGAACAA AGTCAAAGGA 1440
AGCGGCAATT TAGATGTGGC TACCGGGTTG AACTACCGCT ATAAGCATTC TAAATATTCT 1500
GTAGGGATTA GCATCCCTTT AATCCAAAGA AAAGCTAGCG TCGTTTCTAG CGGTGGCGAT 1560
TATACGAACT CTTTTGTTTT CAATGAAGGG GCTAGCCACT TTAAGGTGTT TTTCAATTAC 1620
GGTGGGTGTT TT 1632

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1071 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1071

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTGATGAAAA GCATTTTGCT CTTTATGATT TTTGTAGTTT GTCAGTTAGA AGGCAAAAAA 60
TTTTCACAAG ATAATTTTAA GGTGGATTAT AACTACTATT TGCGCAAACA GGATTTGCAC 120
ATCATTAAAA CGCAAAACGA TTTGTCCAAT GCCTGGTATC TCCCTCCACA AAAAGCCCCC 180
AAAGAACATT CTTGGGTGGA TTTTGCTAAA AAATATTTAA ACATGATGGA TTATCTAGGC 240

ACTTATTTTT TGCCTTTTTA TCATAGTTTC ACCCCCATTT TTCAATGGTA CCACCCTAAT 300
ATCAACCCCT ACCAACGCAA TGAGTTTAAG TTCCAAATCA GTTTTAGAGT GCCTGTATTT 360
AGGCATATTC TTTGGACTAA AGGCACGCTT TATCTGGCTT ATACCCAAAC TAACTGGTTT 420
CAAATTTATA ATGACCCTCA ATCCGCCCCC ATGCGAATGA TCAATTTCAT GCCTGAACTC 480
ATCTATGTTT ATCCTATTAA TTTTAAACCT TTTGGGGGTA AAATAGGGAA TTTTTCTGAA 540
ATTTGGATAG GTTGGCAGCA CATTTCTAAT GGTGTGGGGG GTGCGCAATG TTACCAGCCT 600
TTTAATAAAG AAGGTAATCC TGAAAACCAG TTTCCAGGAC AACCTGTAAT CGTTAAAGAT 660
TATAACGGGC AAAAAGATGT GCGCTGGGGG GGGTGTCKTT CGGTGARCSC GGGCAACSCC 720
CTGTGTTTCG TTTTGGTGTG GGAAAAGGGA GGCCTAAAAA TCATGGTCGC TTATTGGCCC 780
TATGTCCCTT ATGATCAATC CAACCCTCAA TTGATTGATT ACATGGGGTA TGGTAACGCT 840
AAAATTGATT ACAGGAGAGG GCGCCACCAT TTTGAATTGC AACTTTATGA TATTTTCACG 900
CAATACTGGC GTTATGATCG CTGGCATGGA GCTTTCCGCT TAGGCTATAC CTACCGCATT 960
AACCCTTTTG TGGGGATTTA TGCGCAGTGG TTTAACGGCT ATGGCGATGG CTTGTATGAA 1020
TACGATGTTT TTTCCAATCG TATAGGGGTA GGAATACGCT TGAACCCTTA A 1071

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2028 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...2028

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTGTCTAAAG GTTTGAGTAT CGGTAATAAA ATCATATTGT GCGTGGCGTT GATTGTGATC 60
GTGTGCGTGA GCATTTTAGG GGTGTCCTTA AACAGCAGGG TGAAAGAGAT TTTAAAAGAA 120
AGCGCTCTGC ATTCAATGCA AGATAGTTTG CATTTCAAGG TTAAGGAAGT GCAAAGTGTT 180
TTGGAAAACA CTTATACGAG CATGGGCATT GTCAAAGAAA TGCTCCCTGA AGACACCAAA 240
AGAGAAATCA AAATCCAGTT GTTAAAAAAC TTCATTTTAG CCAATTCGCA TGTCGCTGGG 300
GTGAGCATGT TTTTTAAAGA CAGAGAGGAT TTGAGATTGA CGCTTTTACG AGATAACGAT 360
ACGATCAAGT TGATGGAAAA CCCGTCATTA GGGAGTAACC CTTTAGCGCA AAAAGCGATG 420
AAAAATAAAG AAATTTCTAA AAGCTTGCCT TATTACAGGA AAATGCCTAA CGGGGCGGAA 480
GTTTATGGCG TGGATATTCT TTTACCACTA TTCAAGGAAA ACACGCAAGA AGTGGTGGGG 540
GTTCTGATGA TTTTCTTTTC CATTGACAGC TTCAGTAATG AAATCACTAA AAACAGGAGC 600
GATTTATTTT TAATTGGCGT TAAAGGTAAA GTGCTTTTGA GCGCGAATAA AAGCTTGCAA 660
GACAAATCCA TCACCGAAAT TTATAAAAGC GTGCCTAAAG CCACTAATGA AGTGATGGCT 720
ATTTTAGAAA ATGGCTCTAA AGCGACTTTA GAATACTTGG ATCCCTTTAG CCATAAGGAG 780
AATTTTTTAG CCGTTGAAAC CTTTAAAATG CTAGGCAAAA CAGAAAGTAA AGACAATCTT 840
AATTGGATGA TCGCTTTGAT CATTGAAAAA GACAAGGTCT ATGAGCAAGT GGGATCGGTG 900
CGTTTTGTGG TGGTTGCAGC GAGTGCTATC ATGGTGTTAG CCTTAATCAT AGCGATCACT 960
CTTTTAATGC GAGCGATCGT GAGCAATCGT TTGGAAGTCG TTTCTAGCAC CTTGTCTCAT 1020

TTCTTTAAAT TATTGAACAA TCAAGCCCAT TCTAGCGACA TTAAATTGGT TGAAGCGCGA 1080
TCTAATGACG AATTAGGGCG CATGCAAACA GCGATCAATA AAAATATCTT GCAAACCCAA 1140
AAAACCATGC AAGAAGACAG GCAAGCCGTC CAAGACACCA TTAAAGTGGT TTCAGACGTG 1200
AAAGCGGGGA ATTTTGCGGT GCGCATCACG GCTGAACCCG CAAGCCCTGA TTTGAAAGAA 1260
TTGAGAGACG CGCTAAATGG GATCATGGAT TATTTGCAAG AAAGCGTAGG GACTCACATG 1320
CCAAGCATTT TCAAAATCTT TGAAAGCTAT TCTGGCTTGG ATTTTAGAGG GCGGATCCAA 1380
AACGCTTCGG GTAGGGTGGA ATTGGTTACT AACGCTTTAG GGCAAGAAAT CCAAAAAATG 1440
CTAGAAACTT CGTCTAATTT TGCCAAAGAT CTAGCGAACG ATAGCGCGAA TTTAAAAGAA 1500
TGCGTGCAAA ATTTAGAAAA GGCTTCAAAC TCCCAACACA AAAGCCTGAT GGAAACTTCC 1560
AAAACGATAG AAAATATCAC CACTTCCATT CAAGGCGTGA GCTCTCAAAG TGAAGCCATG 1620
ATTGAACAAG GGAAAGACAT TAAAAGCATT GTAGAAATCA TTAGAGATAT TGCCGATCAA 1680
ACGAATCTAT TAGCCCTAAA CGCTGCTATT GAAGCCGCAC GAGCCGGCGA GCATGGCAGA 1740
GGCTTTGCGG TGGTGGCTGA TGAGGTGAGG AAGCTCGCTG AAAGGACGCA AAAATCCCTC 1800
AGTGAGATTG AAGCCAATAT TAATATTCTC GTTCAAAGCA TTTCAGACAC GAGCGAAAGC 1860
ATTAAAAACC AGGTTAAAGA AGTAGAAGAG ATCAACGCTT CTATTGAAGC CTTAAGATCG 1920
GTTACTGAGG GCAATCTAAA AATCGCTAGC GATTCTTTAG AAATCAGTCA AGAAATTGAC 1980
AAAGTCTCTA ACGATATTTT AGAAGATGTG AATAAAAAGC AGTTTTAA 2028

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 816 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...816

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ATGAACATAT TCAAGCGTAT TATTTGCGTA ACCGCTATTG TTTTAGGTTT TTTTAACCTT 60
TTAGACGCCA AACACCACAA AGAAAAAAAA GAAGACCACA AAATCACTCG TGAGCTTAAA 120
GTGGGCGCTA ACCCTGTGCC GCATGCGCAA ATCTTGCAAT CAGTTGTGGA TGATTTGAAA 180
GAGAAAGGGA TCAAATTAGT GATCGTGTCT TTTACGGATT ATGTGTTGCC TAATTTAGCG 240
CTCAATGACG GCTCTTTAGA CGCGAATTAC TTCCAGCACC GCCCTTATTT GGATCGGTTT 300
AATTTGGACA GAAAAATGCA CCTTGTTGGT TTGGCCAATA TCCATGTGGA GCCTTTAAGA 360
TTTTATTCTC AAAAAATCAC AGACATTAAA AACCTTAAAA AAGGCTCAGT GATTGCTGTG 420
CCAAATGATC CGGCCAATCA AGGCAGGGCG TTGATTTTAC TCCATAAACA AGGCCTTATC 480
GCTCTCAAAG ACCCAAGCAA TCTATACGCT ACGGAGTTTG ATATTGTCAA AAATCCTTAC 540
AACATCAAAA TCAAACCCCT AGAAGCTGCG TTATTGCCTA AGGTTTTAGG GGATGTGGAT 600
GGGGCTATCA TAACAGGGAA TTATGCCTTG CAAGCAAAAC TCACCGGAGC CTTATTTTCA 660
GAAGATAAGG ACTCGCCTTA TGCTAATCTT GTAGCCTCTC GTGAGGATAA TGCGCAAGAT 720
GAAGCGATAA AAGCGTTGAT TGAAGCCTTA CAGAGCGAAA AGACCAGGAA ATTCATTTTG 780
GATACCTATA AGGGGGCGAT TATCCCGGCT TTTTAA 816

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 486 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...486

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
ATGTTTTTTA AAACTTATCA AAAATTACTG GGCGCGAGCT GTTTGGCGCT GTATTTAGTG     60
GGCTGTGGGA ATGGTGGTGG CGGTGAATCG CCGGTTGAGA TGATTGCAAA TAGCGAGGGT    120
ACGTTTCAAA TCGACTCCAA AGCAGATAGC ATTACTATTC AAGGCGTGAA GCTTAATAGA    180
GGTAATTGTG CTGTCAATTT TGTTCCAGTA AGTGAGACGT TTCAAATGGG TGTTTTAAGT    240
CAAGTTACTC CAATCTCTAT ACAGGATTTT AAAGATATGG CAAGCACTTA TAAGATATTT    300
GATCAAAAGA AAGGGTTGGC AAACATAGCA AATAAAATTT CTCAATTAGA GCAAAAGGGT    360
GTGATGATGG AACCTCAAAC CCTTAATTTT GGAGAAAGTT TAAAAGGCAT TTCTCAAGGG    420
TGCAATATTA TAGAGGCAGA AATACAAACC GACAAAGGCG CTTGGACTTT TAACTTTGAT    480
AAATAA                                                               486
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1014 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1014

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATGATTAGAT TAAAAGGTTT GAATAAAACT TTAAAAACAA GCTTATTAGC TGGGGTTTTA 60
CTAGGTGCTA CTGCTCCCTT AATGGCAAAG CCTTTATTAA GCGATGAAGA CTTATTGAAA 120
CGAGTAAAAC TACACAATAT CAAAGAAGAT ACGCTGACTA GCTGTAATGC TAAGGTGGAC 180
GGCTCTCAAT ACTTGAATAG TGGTTGGAAT TTATCTAAAG AATTTCCGCA AGAATATAGA 240
GAAAAGATTT TTGAATGCGT AGAAGAAGAA AAACATAAAC AAGCCCTTAA TTTAATCAAT 300
AAAGAAGACA CTGAAGATAA AGAAGAACTT GCAAAAAAAA TCAAAGAAAT TAAAGAAAAA 360
GCTAAAGTTT TAAGGCAAAA ATTTATGGCT TTTGAAATGA AAGAACACTC TAAAGAATTC 420
CCAAATAAAA AGCAACTTCA AACCATGCTT GAGAACGCTT TTGATAATGG AGCTGAAAGT 480
TTTATTGATG ATTGGCACGA ACGCTTTGGG GGTATAAGTA GAGAGAATAC TTATAAAGCA 540
CTTGGCATTA AAGAATATAG TGATGAAGGA AAGATATTAG CCTTTGGCGA AAGAAGTTAT 600
ATTAGACAAT ATAAAAAAGA TTTTGAAGAA AGCACTTATG ATACTAGACA AACCTTATCT 660
GCTATGGCTA ATATGAGTGG CGAAAACGAT TATAAAATTA CTTGGTTAAA ACCCAAATAT 720
CAGCTCCATA GTTCAAATAA TATTAAACCC TTAATGTCAA ACACAGAGTT GTTAAATATG 780
ATAGAGCTAA CCAATATCAA AAAAGAATAT GTTATGGGCT GTAATATGGA AATAGATGGT 840
TCTAAATATC CCATTCATAA AGATTGGGGA TTTTTTGGTA AGGCAAAAGT CCCAGAAACT 900
TGGAGAAATA AGATTTGGGA ATGTATTAAG AATAAAGTAA AGTCCTATGA CAACACTACC 960
GCTGAAATAG GAATAGTTTG GAAAAAAAAT ACTTATTCTA TCTCTCATCA CTAA 1014

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1251 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1251

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATGAAAAAAT TAGTTTTTAG CATGCTTTTA TGTTGTAAAA GCGTGTTTGC AGAGGGGGAA 60
ACTCCTTTGA TTGTCAATGA CCCAGAAACC CATGTAAGTC AAGCCACTAT CATAGGCAAA 120
ATGGTAGATA GTATCAAAAG ATACGAAGAG ATTATTTCTA AGGCTCAAGC TCAAGTCAAT 180
CAGTTACAAA AAGTCAATAA CATGATAAAT ACGACTAATT CTTTGATTAG TAGTAGTGCT 240
ATCACTTTAG CCAATCCTAT GCAAGTTTTA CAAAACGCTC AGTATCAAAT AGAGAGCATT 300
AGATACAACT ATGAGAATTT AAAGCAAAGC ATAGAAAATT GGAACGCACA AAATTTGTTA 360
AGAAACAAAT ACTTACAGCA ACAATGCCCT TGGCTTAATG TCAATGCTCT TACTAACAAT 420
AAGATTGTCA ATCTTAAAGA TCTCAATAAC CTAATCACCA AAAATGGCGA ACAAACCCAA 480
ACCGCAAGAG ATGTGCAAAA TCTCATTCAG TCCATTAGTG GCAGTGGCTA TGGAAACATG 540
CAATCACTTG CTGGGGAATT GAGTGGTAGA GCGTGGGGGG AAATGTTGTG TAAAATGGTA 600
AACGATAGTA ATTATGAAAG CGAGCAAGCT CTTTTAGCAA CAGGCAATAA CCCAGAAGAG 660
CAAAAACGAA GATTTTTGCT TAGAGTAAAG AAAAAGGTTA ATGATAATAA GCAGTTAAAA 720

GATAAACTTG ACCCATTTCT AAAAAGACTT GATGTCCTAC AAACTGAGTT TGGTGTAACT 780
GACCCTACAG CTAACCATAA TAAGCAAGGG ATACATTATT GCACAGAAAA TAAAGAGACA 840
GGTAAATGCG ACCCTATTAA AAATGTATTT AGGACAACTC GCTTAGATAA CGAATTAGAA 900
CAAGAAATCC AAACGCTCAC ACTTGATTTA ATCAAAGCCT CCAATAAAGA CGCTCAAAGC 960
CAAGCCTACG CAAATTTCAA TCAAAGGATT AAATTACTTA CTCTAAAATA TTTAAAAGAA 1020
ATTACCAATC AAATGCTCTT TTTAAATCAA ACAATGGCAA TGCAAAGCGA GATTATGACA 1080
GATGATTATT TTAGGCAAAA TAATGATGGC TTTGGGGAAA AAGAAAACCA TATAGACAAA 1140
CAATTAACGC AAAAAAGAAT AAACGAAAGA GAAAGAGCTA GAATATACTT TCAAAACCCT 1200
AATGTTAAAT TTGACCAATT TGGCTTTCCC ATTTTTAGTA TATGGGATTA A 1251

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1131 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1131

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GTGAATAAGT GGATTAAAGG GGCGGTTGTT TTTGTAGGGG GTTTTGCAAC GATTACAACC 60
TTTTCTTTAA TCTACCACCA AAAGCCAAAA GCCCCCCTAA ATAACCAGCC TAGCCTTTTG 120
AATGACGATG AGGTGAAATA CCCCTTACAA GACTACACTT TCACTCAAAA CCCACAGCCA 180
ACTAACACGG AAAGCTCCAA AGACGCTACC ATCAAAGCCT TACAAGAACA GCTCAAAGCC 240
GCTTTAAAAG CCCTAAACTC CAAAGAAATG AATTATTCCA AAGAAGAGAC TTTTACTAGC 300
CCTCCCATGG ATCCAAAAAC AACCCCCCCT AAAAAAGACT TTTCTCCAAA ACAATTAGAT 360
TTACTGGCCT CTCGCATCAC CCCTTTCAAG CAAAGCCCTA AAAATTACGA AGAAAACCTG 420
ATTTTCCCTG TGGATAACCC TAATGGCATT GATAGTTTCA CTAACCTTAA AGAAAAAGAC 480
ATCGCCACTA ATGAAAACAA GCTTTTACGC ACCATTACAG CTGACAAAAT GATACCCGCT 540
TTTTTGATTA CGCCCATTTC TAGCCAGATC GCTGGTAAAG TGATTGCGCA AGTGGAGAGC 600
GATATTTTTG CAAGCATGGG CAAAGCCGTC TTAATCCCCA AAGGCTCTAA AGTCATAGGC 660
TATTACAGCA ACAATAACAA AATGGGCGAA TACCGCTTGG ATATTGTATG GAGTCGAATC 720
ATCACTCCCC ATGGCATTAA TATCATGCTC ACTAACGCTA AAGGGGCGGA CATTAAAGGC 780
TATAACGGCT TAGTGGGGGA ATTGATTGAA AGGAATTTCC AACGCTATGG CGTGCCGTTA 840
CTGCTTTCTA CGCTCACTAA CGGCCTATTG ATTGGGATCA CTTCGGCTTT AAACAACAGA 900
GGCAATAAAG AAGAGGTGAC TAATTTCTTT GGGGATTATC TTTTATTGCA ATTGATGAGG 960
CAAAGCGGCA TGGGGATCAA TCAAGTGGTC AATCAAATTT TAAGAGACAA GAGCAAGATC 1020
GCCCCCATTG TGGTGATTAG AGAGGGGAGT AGGGTCTTCA TTTCGCCCAA TACTGACATC 1080
TTCTTCCCTA TACCCAGAGA GAATGAAGTC ATCGCTGAGT TTTTGAAGTG A 1131

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2751 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...2751

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTGGATTTGA GGATCCAATC TAAAGAAGTC AGTCATAATT TAAAGGAATT ATCAAAAACG 60
CTAATCAGCT ATCCTTTTGA AAAACATGTA GAAGCTTTAG GGGAACAATG CAGTAACTTC 120
GTTTCTATTC CCATTAACAA TGACGACTAT TCAAATATTT GCACTTTTGT GAGTGATTTT 180
ATAAATCTTA TAGCTTCTTA CAATTTATTA GAATCATTTT TAGATTTTTA TAAAGATAAA 240
TTAAAATTGA GCGAGCTTGT AACTGAATAT GCCAACGTAA CCAATAATCT GCTTTTCAAA 300
AAATTAATCA AACATTTAAG CGGCAACAAT CAATTGGTTA AAAATTTTTA TCAGTGTATA 360
AGAGAAATTA TAAAATACAA CGCCCCTAAT AAAGAATACA AACCCAATCA ATTTTTTATA 420
ATAGGGAAAG GCAAACAAAA ACAATTAGCA AAAATTTATT CTCATTTAAA AGAACTTAGT 480
GCAAGTGAAA TTAAACCACA AGATATGGAA GACATCTTAA AAAAGCTAGA GGAATTAGAT 540
AAAATTTTTA AAACTACCGA CTTTACAAAA TTCACACCAA AAACTGAAAT TAAGGATATT 600
ATTAAAGAAA TAGACGAAAA ATACCCTATC AATGAAAATT TTAAACGGCA ATTTAATGAG 660
TTTGAATCAA ATATTGAAAA ACATGATGAA ATAAAAAAGG ATTTTGAGCG AAACAAAGAG 720
TCGCTGATCC GAGAAATTGA AAATCACTGC AAAAATGAAT GCAATAGCGA AGAAGAGCCG 780
GAGTATAAGA TTAATGATCT GCTCAAAAAT ATCCAACAAA TATGCAAAAA TTATATAGAA 840
AGTCATGCCG TTAATGATGT GTCTAAAGAT ATTAAATCCA TGATGTGTCA GTTTTATTTG 900
AAACAGATAG ATTTATTAGT CAATTCAGAA ATTGTGCGAT ACAGATACAG CAATCTTTTT 960
GAACCAATAC AAAGATCTTT ATGGGAGAGT ATAAAAATTT TAGATAATGA AAGTGGCATT 1020
TATTTGTTCC CTAAAAATAT TGGTGAAATC AAGGATAAAT TTGAAGCAAA CAAGGAAAAA 1080
TTCAAACAAA GCAAAAATGT TTCTGAGTTC GCAGAATATT GCCGAGAGTG TAACCCCTAT 1140
ACAGCGTTTA ACTTTCATCT AAATATAAAT AATGGTTTAT CTCATCAATT TGAAAAATTC 1200
GTGCCAATCA TGAAAGAATA CAAAGAGCCA AAAATCACAG ATAATGACCT TGAAGCCATA 1260
TCAACCAAAG AGACTGGTCT TGCTAGCCAA TTATCTGGGC ACTGGTTTTT TCAGCTTTCG 1320
TTATTTAATA AAACAAACTT TAATCCTAAT AAAATTTGGA TTCCTTTAGA GTTCAATAAA 1380
AGATCAAAAA TAAAGTTTGA TAAAGATTTA GAAATCTATT TTGATAGTCA TGAATCGTTC 1440
AATATCTCTA AAAAATACTT GCAAGAAATA GATCAAGAAT CACTAAAAAA GATCAAACAA 1500
TCAAAAGATT TTTTTTCAAT TCAAAAAATA GAGAGTAAGC ATGATAATAA CGATATACTG 1560
CAACTTGAAT TTTTTGAGAA TGATACAAGT TTTCTTTTTG CTAAAGGAAG TTTTGCAGAA 1620
ATTTTAGAAT ACAACATGCA ATTAAAAATA GATTCTTTAA TTACAAAAGA ATTTAATAAG 1680
CTTTTAGCGA TCGTTCAAGA TAGTCCCCAA GATAGTTACC AATTAAAAAT TCGTGTCCGA 1740
CATAACAATA AGCTTCCTAG AGAGAAATAT ACGGAACATG AAATAAAACT TGAAGTTTAT 1800
GATTGCAGAA AATCCCACGA TCACAATGAG CCAATCATCT TAAGCCAGCA AAGCACCGGC 1860
TTCCAATGGG CGTTTAATTT CATGTTTGGC TTTCTTTATA ATGTGGGATC ACATTTTAGT 1920
TTTAACCATA ATATTATCTA TGTCATGGAC GAGCCAGCCA CTCATTTGAG CGTGCCAGCC 1980
AGAAAGGAGT TTAGGAAATT TTTAAAAGAA TACGCTCATA AAAATCATGT TACTTTTGTT 2040

TTAGCCACCC ATGACCCCTT TTTAGTGGAT ACGGATCATT TAGATGAAAT AAGGATTGTG 2100
GAAAAGGAAA CAGAAGGCTC TGTAATTAAG AATCACTTTA ACTATCCCCT AAATAATGCA 2160
AGCAAAGACT CCGACGCTTT GGACAAAATC AAACGCTCTT TAGGAGTGGG CCAGCATGTT 2220
TTTCATAACC CCCAAAAACA CCGAATCATT TTTGTAGAAG GCATCACGGA TTATTGTTAT 2280
TTGAGCGCTT TTAAATTGTA TTTGCGTTAC AAAGAATACA AGGACAACCC CATTCCTTTC 2340
ACTTTCTTAC CCATTTCAGG GCTTAAAAAC GATTCAAACG ATATGAAAGA AACCATTGAA 2400
AAACTTTGCG AGTTAGACAA TCACCCTATT GTTTTGACAG ACGATGACAG AAAATGCGTT 2460
TTTAACCAAC AAGCAACGAG CGAACGATTT AAAAGAGCTA ATGAAGAAAT GCATGATCCC 2520
ATCACCATCC TACAACTCTC AGACTGCGAT AGGCATTTCA AACAAATTGA AGATTGTTTC 2580
AGCGCAAACG ATAGAAACAA ATACGCTAAA AATAAGCAAA TGGAATTGAG CATGGCTTTT 2640
AAAACAAGGC TTTTGTATGG CGGAGAAGAT GCGATAGAAA AACAAACAAA AAGAAATTTT 2700
TTAAAATTAT TCAAATGGAT TGCATGGGCT ACAAACTTGA TCAAAAACTA A 2751

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 531 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...531

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATGACTGCAA TGATGCGTTA TTTTCACATC TATGCGACCA CTTTTTTCTT CCCTTTGGCG 60
CTTCTTTTTG CGGTTAGTGG GCTTTCATTG CTCTTTAAAG CGCGCCAAGA CACTGGCGCT 120
AAGATCAAAG AATGGGTTTT AGAAAAATCC TTAAAAAAAG AAGAACGATT GGACTTTTTA 180
AAAGGCTTTA TAAAAGAAAA CCATATCGCT ATGCCTAAAA AGATAGAGCC TAGAGAGTAT 240
AGGGGAGCGT TAGTCATTGG CACGCCTTTG TATGAAATCA ACCTTGAAAC TAAAGGCACT 300
CAAACGAAAA TCAAGACCAT TGAAAGGGGC TTTTTAGGCG CGCTCATCAT GCTGCATAAG 360
GCTAAGGTGG GCATCGTGTT TCAGGCGCTT TTAGGGATTT TTTGCGTGTT TTTATTGTTG 420
TTTTACTTGA GCGCGTTTTT AATGGTGGCT TTTAAAGACA CTAAACGCAT GTTTATAAGC 480
GTTTTAATAG GGAGCGTGGT GTTCTTTGGA GCGATCTATT GGTCTTTGTA G 531

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 669 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...669

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATGTTTAAAA ACGCTTTAAA TATACAAGAT TTTTCATTTA AAAATCATAC TAGTACAGCC 60
ATTATTGGCA CAAATGGTGC TGGAAAATCA ACGCTTATCA ACACTATTCT AGGCATTAGA 120
TCAGACTATA ATTTTAAAGC ACAAAACAAT AATATTCCAT ACCACGACAA TGTTATACCA 180
CAACGCAAGC AATTGGGAGT TGTCTCTAAC CTATTCAACT ACCCACCTGG ATTAAACGCA 240
AACGACCTTT TTAAATTCTA TCAATTTTTT CACAAAAACT GCACTCTAGA TTTGTTTGAA 300
AAAAATCTTT TAAATAAAAC CTACGAACAC CTAAGCGACG GACAAAAACA GCGCTTAAAA 360
ATTGACTTAG CTCTTAGCCA TCACCCACAA TTAGTTATTA TGGATGAACC AGAAACCAGT 420
TTAGAGCAAA ACGCTCTTAT AAGACTATCA AATCTCATAA GCTTGCGCAA CACCCAACAA 480
CTTACAAGTA TCATCGCCAC TCATGATCCT ATTGTCTTAG ATAGTTGCGA ATGGGTATTG 540
CTCCTTAAGA ATGGCAACAT TGCTCAATAC AAACCTTTAA ATTCTATATT AAAATCTGTA 600
GCTAAAACTT TTAACTTTAA AGAAAAACCA ACCACAAAAG ACTTATTAGC GTTACTAAAG 660
GATATTTAA 669

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1221 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1221

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATGTATGCGG CTCATCCTAT TAAACCCATA AAAGCCCCTA AACTCAAATC TCAATTTTTA 60
AGGCGTGTGT TTGTGGGCGC GTCCATTAGG CGCTGGAATG ACCAAGCATG CCCTTTGGAA 120
TTTGTGGAAT TAGACAAGCA AGCCCATAAA GCGATGATTG CGTATCTGCT CGCTAAAGAT 180
TTAAAAGATA GGGGTAAAGA TTTAGATTTA GATCTTTTAA TCAAATATTT TTGCTTTGAG 240
TTTTTGGAGC GCTTGGTTTT AACCGATATT AAACCCCCTA TTTTTTACGC CCTCCAACAA 300

ACGCATAGTA AAGAGTTAGC TTCCTATGTT GCGCAAAGTT TGCAAGATGA AATCAGTGCG 360
TATTTTTCTT TAGAGGAACT CAAAGAGTAT TTAAGCCACA GGCCTCAAAT TTTAGAAACT 420
CAAATTTTAG AGAGCGCGCA TTTTTATGCG TCTAAGTGGG AGTTTGATAT TATCTATCAT 480
TTTAACCCCA ACATGTATGG CGTGAAAGAG ATTAAAGATA AAATTGACAA GCAACTCCAC 540
AATAACGATC ATTTGTTTGA AGGGCTTTTT GGGGAAAAAG AAGATTTGAA AAAATTGGTG 600
AGCATGTTTG GGCAGTTGCG TTTCCAAAAG CGCTGGAGCC AAACCCCAAG AGTGCCACAA 660
ACCAGTGTTC TAGGGCATAC TTTATGCGTG GCGATTATGG GGTATTTATT GAGTTTTGAC 720
TTGAAAGCTT GTAAAAGCAT GCGGATCAAT CATTTTTTGG GCGGGCTTTT CCATGATTTA 780
CCCGAAATTT TAACCCGAGA CATTATCACG CCCATCAAAC AAAGCGTTGC AGGGCTTGAT 840
CATTGCATTA AAGAGATTGA AAAAAAGGAA ATGCAAAACA AAGTCTATTC CTTTGTGTCT 900
TTGGGCGTTC AAGAAGATTT GAAATATTTC ACCGAAAACG AGTTTAAAAA CCGCTACAAA 960
GACAAGTCTC ATCAAATCGT TTTCACTAAA GACGCTGAAG AATTATTCAC GCTTTATAAT 1020
AGCGATGAAT ATCTTGGGGT TTGCGGGGAG CTTTTGAAGG TGTGCGATCA TTTGAGCGCG 1080
TTTTTAGAAG CCCAAATCTC TCTTTCTCAT GGCATTTCTA GCTACGATTT AATCCAAGGA 1140
GCTAAAAACC TTTTAGAATT GCGATCCCAA ACGGAACTGC TTGATTTGGA TTTAGGGAAA 1200
TTGTTTAGAG ATTTTAAGTA A 1221

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1008 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1008

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTGTTGTGGG TGCTATATTT TTTAACCAGT TTATTTATTT GCTCTTTGAT TGTTTTGTGG 60
TCTAAAAAAT CCATGCTCTT TGTGGATAAC GCTAATAAAA TCCAAGGCTT CCATCATGCA 120
AGAACCCCAC GAGCCGGGGG GCTTGGGATC TTTCTTTCTT TTGCGTTGGC TTGTTATCTT 180
GAACCTTTTG AGATGCCTTT TAAGGGGCCT TTTGTTTTCT TAGGGCTATC GCTAGTGTTT 240
TTGAGCGGTT TTTTAGAAGA CATTAACCTT TCATTAAGCC CCAAAATACG CCTTATTTTG 300
CAAGCTGTAG GGGTCGTTTG CATCATTTCA TCAACGCCTT TAGTGGTGAG CGATTTTTCG 360
CCCCTTTTTA GCTTGCCTTA TTTCATCGCT TTTTTATTCG CTATTTTTAT GCTGGTGGGT 420
ATCAGTAACG CTATTAATAT CATTGACGGG TTTAACGGGC TTGCATCTGG GATTTGCGCG 480
ATCGCGCTTT TAGTCATTCA TTATATAGAC CCTAGCAGTT TGTCTTGTTT GCTCGCTTAC 540
ATGGTGCTTG GGTTTATGGT GTTAAATTTC CCTTCAGGAA AGATTTTTTT AGGCGATGGG 600
GGGGCGTATT TTTTGGGTTT GGTGTGCGGG ATTTCTCTCT TGCATTTGAG TTTGGAGCAA 660
AAAATCAGCG TGTTTTTTGG GCTCAATTTA ATGCTTTATC CGGTCATAGA GGTGCTTTTT 720
AGTATCCTTA GGCGCAAAAT AAAACGCCAG AAAGCCACCA TGCCGGATAA TTTGCATTTG 780
CACACCCTTT TATTTAAATT CTTGCAACAA CGCTCTTTCA ATTACCCTAA CCCTTTATGC 840
GCGTTTATCC TTATTCTATG CAACCTGCCT TTTATTTTAA TAAGCGTTTT GTTTCGCTTG 900

GACGCTTATG CGCTCATTGT GATTAGCCTA GTCTTTATCG CATGCTATTT AATAGGCTAT 960
GCTTATTTGA ATAGGCAAGT TTGCGCTTTA GAAAAGCGGG CGTTTTAA 1008

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 291 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...291

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ATGAAAAAGG TTATTGTGGC TTTAGGCGTT TTGGCGTTCG CAAATGTTTT AATGGCAACC 60
GATGTTAAGG CTCTTGTAAA AGGTTGTGCC GCTTGCCATG GGGTTAAGTT TGAAAAGAAA 120
GCTTTAGGTA AAAGCAAAAT CGTTAACATG ATGAGCGAAA AAGAGATTGA AGAGGATCTT 180
ATGGCTTTTA AAAGCGGTGC CAACAAGAAT CCTGTCATGA CCGCGCAAGC TAAAAAATTA 240
AGCGATGAAG ACATCAAAGC TTTAGCCAAA TACATCCCCA CTCTCAAATA A 291

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 471 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...471

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ATGCGAGATT TCAATAACAT TCAAATCACA CGCTTAAAAG TGCGTCAAAA TGCCGTTTTT 60
GAAAAACTGG ATCTGGAGTT TAAAGATGGC TTGAGCGCGA TTAGTGGGGC TAGTGGGGTG 120
GGGAAAAGCG TCCTTATTGC GAGCCTTTTA GGGGCGTTTG GGCTTAAAGA GAGCAACGCT 180
TCAAACATTG AAGTGGAATT GATCGCGCCT TTTTTAGACA CGGAAGAATA CGGCATTTTT 240
AGAGAAGATG AGCATGAACC CTTAGTTATT AGCGTGATTA AAAAAGAAAA AACACGCTAT 300
TTTTTAAACC AAACAAGCCT ATCTAAAAAC ACGCTCAAAG CGTTATTAAA GGGGCTTATT 360
AAACGCTTAT CTAACGACAG ATTCAGCCAG AATGAACTCA ACGATATTTT AATGCTCTCC 420
TTATTAGATG GCTATATCCA AAATAAAAAT AGGCGTTTAG CCCCCTTTTA G 471

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 357 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...357

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTGATGCTAA TGGCAATTTT TACCCCTTAT ATTCTTATTT TGAAAATGAT GAAAAAGTCT 60
ATGAGTTTAT TCGCCAATAT GGGGTTGGAG CAAATTTTTT GCAACAGAGA CATTAAAGAT 120
TTAAATGATT TTGTTTTTGG TATAGAAGTG GGGCTTGATA GCAATGCGAG AAAAAATCGT 180
AGCAGAAAGG CTATGGAAAA TCATCTTATC GGTCTTTTTG TCCAAGCTCA ATTAAATTTT 240
AAAGAACAAG TAGATATTAG AGAATTTGAG GATTTACGCC AGGCTTTTGG AAATGATACT 300
AAAAAATTTG ATTTTGTTAT TTTTAGCAAA GAGAAAACTT ATTTTCATAG AAGCTAA 357

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1068 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1068

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATGAATATCA AAATTTTAAA AATATTAGTT GGAGGGTTAT TTTTTTTGAG CTTGAACGCC 60
CATTTATGGG GGAAACAAGA CAATAGCTTT TTAGGGATTG GTGAAAGAGC CTATAAAAGC 120
GGGAATTATT CTAAAGCGGC GTCTTATTTT AAAAAAGCAT GCAACGATGG GGTGAGTGAA 180
GGCTGCACGC AATTAGGAAT CATTTATGAA AACGGGCAAG GCACTAGAAT AGATTATAAA 240
AAAGCCCTAG AATATTATAA AACCGCATGC CAGGCTGATG ATAGGGAAGG GTGTTTTGGC 300
TTAGGGGGGC TTTATGATGA GGGTTTAGGC ACGGCTCAAA ATTATCAAGA AGCCATTGAC 360
GCTTACGCTA AGGCATGCGT TTTAAAACAC CCTGAGAGTT GCTACAATTT AGGCATCATT 420
TATGATAGAA AAATCAAAGG CAATGCCGCT CAAGCGGTTA CTTACTATCA AAAAAGCTGT 480
AATTTTGATA TGGCTAAGGG GTGTTATATT TTAGGCACTG CCTATGAAAA AGGCTTTTTA 540
GAAGTCAAAC AGAGCAACCA TAAAGCCGTT ATCTATTATT TGAAAGCGTG CCGATTGAAT 600
GAGGGGCAGG CTTGCCGAGC GTTAGGGAGT TTGTTTGAAA ATGGCGATGC AGGGCTTGAT 660
GAAGATTTTG AAGTGGCGTT TGATTATTTG CAAAAAGCTT GCGCTTTAAA CAATTCTGGT 720
GGTTGCGCGA GTTTAGGCTC TATGTATATG TTGGGCAGGT ATGTTAAAAA AGACCCCCAA 780
AAGGCTTTTA ACTATTTCAA GCAAGCATGC GATATGGGGA GCGCGGTGAG TTGCTCTAGG 840
ATGGGCTTTA TGTATTCGCA AGGGGACACT GTTTCAAAAG ACTTGAGGAA AGCCCTTGAT 900
AATTATGAAA GAGGTTGCGA TATGGGCGAT GAAGTGGGTT GCTTCGCTCT AGCGGGCATG 960
TATTACAACA TGAAAGATAA AGAAAACGCC ATAATGATTT ATGACAAGGG CTGTAAATTG 1020
GGCATGAAAC AGGCATGCGA AAATCTCACC AAACTCAGGG GGTATTAG 1068

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 582 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...582

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ATGAAAGAAA AAAACTTTTG GCCTTTAGGA ATCATGAGCG TGCTTATTTT TGGGCTTGGG 60
ATCGTGGTGT TTTTAGTGGT GTTTGCCCTA AAAAATTCGC CTAAAAATGA TTTAGTGTAT 120
TTCAAGGGTC ATAACGAAGT GGATTTAAAC TTTAACGCCA TGCTTAAAAC TTATGAAAAC 180
TTTAAATCCA ATTATCGTTT TTCAGTGGGT TTAAAGCCTC TTACCGAAAG CCCTAAAACC 240
CCCATTTTGC CCTATTTTTC TAAAGGCACG CATGGGGATA AAAAAATCCA AGAAAACCTT 300
TTAAACAACG CTTTGATTTT AGAAAAGTCC AACACGCTTT ATGCACAATT GCAACCGCTC 360

AAACCCGCTT TAGATTCGCC AAATATTCAA GTGTATTTAG CGTTCTATCC CAGCCAATCC 420
CAGCCCAGAT TATTAGGAAC GCTTGATTGT AAAAACGCAT GCGAACCTTT AAAATTTGAT 480
TTGTTAGAGG GCGATAAAGT GGGGCGCTAT AAGATCCTTT TTAAATTTGT TTTTAAAAAT 540
AAAGAAGAAT TGATTTTGGA GCAACTGGCT TTTTTTAAGT AG 582

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 870 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...870

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TTGGGTATCA ATATGTGTTC TAAAAAAATA AGAAATCTCA TTTTATGCTT TGGTTTTATT 60
TTAAGCTTGT GCGCTGAAGA AAATATCACC AAAGAAAACA TGACTGAAAC GAACACGACT 120
GAAGAAAACA CCCCTAAAGA CGCTCCCATT CTTTTGGAAG AAAAACGCGC CCAAACTCTA 180
GAGCTTAAAG AAGAAAATGA AGTGGCAAAA AAGATTGATG AAAAAAGCCT GCTTGAAGAA 240
ATCCATAAGA AAAAACGCCA GCTTTACATG CTCAAAGGGG AATTGCATGA AAAGAATGAA 300
TCCATCTTAT TCCAACAAAT GGCTAAAAAT AAGAGCGGCT TTTTTATAGG CGTGATCCTT 360
GGCGATATAG GGATTAACGC TAATCCTTAT GAGAAGTTTG AACTTTTAAG CAATATTCAA 420
GCTTCTCCCT TGCTGTATGG TTTAAGGAGC GGGTATCAAA AGTATTTCGC TAACGGGATT 480
AGCGCCTTAC GCTTTTATGG GGAATATTTA GGGGGGGCGA TGAAAGGGTT TAAAAGCGAT 540
TCTTTAGCTT CTTATCAAAC CGCAAGCTTG AATATTGATC TGTTGATGGA TAAGCCTATT 600
GACAAAGAAA AAAGGTTTGC GTTAGGGATA TTTGGAGGCG TTGGAGTGGG GTGGAATGGG 660
ATGTATCAAA ATTTAAAAGA GATTAGAGGG TATTCACAGC CTAACGCCTT TGGGTTGGTG 720
TTAAATTTAG GGGTGAGCAT GACGCTCAAC CTCAAACACC GCTTTGAATT AGCCCTAAAA 780
ATGCCTCCCT TAAAAGAAAC TTCGCAAACC TTTTTATATT ATTTTAAAAG CACTAATATT 840
TATTATATTA GTTACAACTA TTTATTGTAA 870

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2007 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...2007

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

ATGAGAAAAC TATTCATCCC ACTTTTATTA TTCAGCGCTT TAGAAGCGAA CGAGAAAAAC 60
GGCTTTTTCA TAGAAGCCGG CTTTGAAACT GGGCTATTAG AAGGCACACA AACGCAAGAA 120
AAAAGACACA CCACCACAAA AAACACTTAC GCAACTTACA ATTATTTACC CACAGACACG 180
ATTTTAAAAA GAGCGGCTAA TTTATTCACC AATGCCGAAG CGATTTCAAA ATTAAAATTC 240
TCATCTTTAT CCCCTGTTAG AGTGTTGTAT ATGTATAATG GTCAATTAAC TATAGAAAAC 300
TTCTTGCCTT ATAATTTAAA TAATGTTAAG CTTAGTTTTA CAGACGCTCA AGGCAACACG 360
ATTGATCTAG GCGTGATAGA GACCATCCCC AAACACTCTA AGATTGTTTT ACCCGGGGAG 420
GCGTTTGATA GTTTAAAAGA GGCGTTTGAT AAAATTGACC CCTATACTTT ATTTCTTCCA 480
AAATTTGAAG CCACTAGCAC TTCTATTTCT GATACTAACA CGCAGAGGGT GTTTGAAACG 540
CTCAATAACA TTAAAACAAA TCTTATAATG AAATATAGTA ATGAAAATCC AAACAATTTC 600
AACACTTGTC CTTACAATAA TAATGGTAAT ACAAAAAATG ATTGTTGGCA AAATTTCACC 660
CCACAAACCG CAGAAGAATT CACCAATTTA ATGTTGAACA TGATCGCTGT CTTAGACTCC 720
CAATCTTGGG GCGATGCGAT CTTAAACGCT CCTTTTGAAT TCACTAACAG CTCAACAGAT 780
TGCGATAGCG ATCCTTCAAA ATGCGTAAAT CCCGGAGTAA ATGGGCGTGT TGATACTAAA 840
GTCGATCAAC AATATATACT CAACAAACAA GGTATTATTA ATAATTTTAG AAAAAAAATA 900
GAAATTGATG CGGTTGTTTT AAAAAATTCA GGGGTTGTAG GGTTAGCCAA TGGATATGGC 960
AATGATGGTG AATATGGCAC ATTAGGGGTA GAAGCCTATG CTTTAGATCC TAAAAAACTC 1020
TTTGGCAACG ACCTTAAGAC TATCAATTTA GAAGATTTAA GAACCATCTT GCATGAATTC 1080
AGCCACACTA AAGGCTATGG GCATAACGGG AATATGACCT ATCAAAGAGT GCCGGTAACG 1140
AAAGATGGTC AAGTGGAAAA GGATAGTAAT GGCAAGCCAA AAGATTCTGA TGGCCTCCCC 1200
TATAATGTGT GTTCGCTTTA TGGGGGATCC AATCAGCCCG CTTTCCCTAG CAACTACCCT 1260
AATTCCATCT ATCACAATTG TGCGGATGTC CCGGCTGGCT TTTTAGGGGT AACAGCAGCG 1320
GTTTGGCAGC AGCTCATCAA TCAAAACGCC TTGCCGATCA ACTACGCTAA CTTGGGGAGT 1380
CAAACAAACT ACAACCTAAA CGCTAGTTTA AACACGCAAG ATTTAGCCAA TTCCATGCTC 1440
AGCACCATCC AAAAAACCTT TGTAACTTCT AGCGTTACCA ACCACCATTT TTCAAACGCA 1500
TCGCAAAGTT TTAGAAGCCC TATTTTAGGG GTTAACGCTA AAATAGGCTA TCAAAACTAC 1560
TTTAATGATT TCATAGGGTT GGCTTATTAT GGCATCATCA AATACAATTA CGCTAAAGCT 1620
GTTAATCAAA AAGTCCAGCA ATTGAGCTAT GGTGGGGGGA TAGATTTGTT ATTGGATTTC 1680
ATCACCACTT ACTCCAATAA AAATAGCCCT ACAGGCATTC AAACCAAAAG GAATTTTTCT 1740
TCATCTTTTG GTATCTTTGG GGGGTTAAGG GGCTTGTATA ACAGCTATTA TGTGTTGAAC 1800
AAAGTCAAAG GAAGCGGCAA TTTAGATGTG GCTACCGGGT TGAACTACCG CTATAAGCAT 1860
TCTAAATATT CTGTAGGGAT TAGCATCCCT TTAATCCAAA GAAAAGCTAG CGTCGTTTCT 1920
AGCGGTGGCG ATTATACGAA CTCTTTTGTT TTCAATGAAG GGGCTAGCCA CTTTAAGGTG 1980
TTTTTCAATT ACGGGTGGGT GTTTTAG 2007

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 192 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...192

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ATGAATACAG AAATTTTAAC CATCATGTTA GTTGTCTCCG TGCTTATGGG ATTGGTAGGC 60
TTAATAGCGT TTTTATGGGG GGTTAAAAGC GGTCAGTTTG ACGATGAAAA ACGCATGCTT 120
GAAAGCGTGT TGTATGACAG CGCGAGCGAC TTGAACGAAG CGATTTTACA AGAAAAACGC 180
CAAAAGAATT AA 192

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1221 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1221

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

ATGGTATTTT TTCATAAGAA AATTATTTTA AATTTTATCT ATTCTTTAAT GGTTGCTTTT 60
TTATTCCATT TATCCTATGG GGTTCTTTTA AAAGCCGATG GAATGGCTAA AAAGCAAACT 120
CTTTTAGTGG GTGAAAGGCT TGTGTGGGAT AAGCTCACGC TGTTAGGGTT TTTAGAAAAA 180
AACCATATCC CCCAAAAACT CTACTACAAT TTGAGCTCTC AAGATAAAGA ATTGAGTGCT 240
GAAATCCAAA GCAATGTTAC CTACTACACT TTAAGAGATG CAAATAACAC GCTCATTCAA 300
GCCCTTATCC CTATTAGCCA GGATTTGCAA ATCCATATTT ACAAAAAAGG AGAGGATTAT 360
TTTTTAGACT TTATCCCCAT TGTTTTCACT CGTAAAGAAA GAACCCTCCT TCTTTCCTTA 420
CAAACTTCGC CCTATCAAGA TATTGTCAAA GCCACCAATG ACCCCCTTTT AGCCAACCAA 480
TTGATGAACG CGTATAAAAA AAGCGTGCCT TTTAAACGCC TAGTGAAAAA CGATAAAATC 540
GCTATCGTTT ATACAAGGGA TTATCGTGTG GGGCAAGCGT TTGGCCAGCC GACCATCAAA 600
ATGGCGATGG TTAGCTCTCG TTTGCACCAA TACTATCTTT TTTCCCATTC AAACGGGCGT 660
TATTACGATT CAAAAGCGCA AGAAGTGGCA GGGTTTTTAC TAGAAACCCC GGTGAAATAC 720

ACCCGCATTT CTTCGCCTTT TTCGTATGGG AGGTTCCATC CTGTTTTAAA AGTTAAACGG 780
CCTCATTACG GCGTGGATTA TGCGGCTAAA CATGGCAGTT TGATCCATTC TGCTTCAGAC 840
GGCCGTGTGG GTTTTATAGG GGTTAAGGCG GGTTATGGGA AGGTGGTTGA AATCCATTTG 900
AATGAATTGC GCTTGGTGTA TGCTCACATG AGCGCGTTCG CTAACGGATT AAAAAAAGGC 960
TCGTTCGTTA AAAAAGGGCA AATCATAGGA AGAGTGGGAA GCACGGGTTT AAGCACCGGG 1020
CCGCATTTGC ATTTTGGCGT GTATAAAAAC TCCCGCCCCA TTAATCCTTT AGGCTATATC 1080
CGCACCGCTA AAAGCAAGCT GCATGGCAAA CAAAGAGAGG TTTTTTTAGA AAAAGCTCAG 1140
TATTCTAAGC AAAAATTAGA AGAACTTTTT AAAACCCATT CTTTTGAAAA AAATTCATTT 1200
TATCTTTTAG AGGGTTTTTA A 1221

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 891 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...891

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TTGTTTTTAG TCAAAAAAAT AGGCGTGGTA ATAATGATTT TAGTCTGCTT TTTAGCTTGC 60
TCGCAAGAGA GCTTTATCAA AATGCAAAAA AAAGCCCAAG AGCAAGAAAA TGACGGCTCT 120
AAACGCCCCA GCTATGTGGA TTCGGATTAT GAAGTCTTTA GCGAAACGAT TTTTTTACAA 180
AACATGGTGT ATCAGCCTAT AGAGGAAAGA AACGCTTTTT TCCAACTGAC TAAAGATGAA 240
GACAATTCTT TTAACCCTGA AAATTCCGTG ATTTTACTGA ATGAGCCAAG CGATAATAGT 300
GAAAAAAACC TACTCTCATA CCCAAACGAT CCCAATAACA ATGAAGACAA CGCTAATAAT 360
AGTCAAAAAA ATCCGTTCCT TTACAAGCCC AAAAGAAAAA CAAAAAACCC AAAACTCATT 420
GAATATTCCC AACAAGATTT CTACCCCCTA AAAAATGGGG ATATTATCAT GAGTAAAGAA 480
GGGGATCAAT GGTTGATAGA AATCCAATCC AAAGCCTTGA AGCGTTTTTT AAAAGATCAA 540
AACGATAAAG ATCGCCAGAT CCAAACTTTC ACTTTTAATG ACACTAAAAC GCAAATCGCG 600
CAAATTAAGG GCAAAATTTC TTCGTATGTT TATACCACCA ATAACGGTAG CTTGAGTTTA 660
AGGCCTTTTT ATGAATCGTT TTTGTTAGAA AAAAAGAGCG ATAATGTTTA TACGATAGAG 720
AATAAGGCTT TAGATACTAT GGAGATTTCA AAGTGTCAAA TGGTGTTAAA AAAGCATTCA 780
ACCGATAAAT TAGACAGCCA GCATAAAGCC ATCAGTATTG ATTTGGATTT TAAAAAAGAG 840
CGCTTTAAGA GCGATACGGA ACTCTTTTTA GAATGTCTTA AGGAAAGTTA G 891

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 747 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...747

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GTGAGCTATG ACAACACCGA TGATTATTAT TTCCCTAGAA ATGGGGTTAT CTTTAGTTCC 60
TATGCGACAA TGTCTGGTTT GCCAAGCTCT GGCACGCTCA ATTCTTGGAA CGGGTTAGGC 120
GGGAATGTCC GTAACACCAA AGTTTATGGT AAATTCGCCG CTTACCACCA TTTGCAAAAA 180
TATTTATTGA TAGATTTGAT CGCTCGTTTT AAAACGCAAG GGGGCTATAT CTTTAGGTAT 240
AACACCGATG ATTACTTGCC CTTAAACTCC ACTTTCTACA TGGGGGGCGT AACCACGGTG 300
AGAGGCTTTA GGAACGGCTC AATCACACCT AAAGATGAGT TTGGCTTGTG GCTTGGAGGC 360
GATGGGATTT TTACCGCTTC TACTGAATTG AGCTATGGGG TGTTAAAAGC GGCTAAAATG 420
CGTTTAGCGT GGTTTTTTGA CTTTGGTTTC TTAACCTTTA AAACCCCAAC TAGGGGGAGT 480
TTCTTCTATA ACGCTCCCAC CACGACGGCG AATTTTAAAG ATTATGGCGT TGTAGGGGCT 540
GGGTTTGAAA GGGCGACTTG GAGGGCTTCT ACAGGCTTAC AGATTGAATG GATTTCGCCC 600
ATGGGGCCTT TGGTGTTGAT TTTCCCTATA GCGTTTTTCA ACCAATGGGG CGATGGCAAT 660
GGCAAAAAAT GTAAAGGGCT GTGCTTTAAC CCTAACATGA ACGATTACAC GCAACATTTT 720
GAATTTTCTA TGGGAACAAG GTTTTAA 747

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1008 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1008

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GTGCAACACT TCAATTTCCT CTATAAAGAT TCTTTATTTT CTATCGCTTT ATTCACTTTC 60

ATTATCGCTC TTGTGATTTT ATTAGAACAG GCTAGAGCGT ATTTCACCCG AAAGAGAAAC 120
AAAAAATTTT TGCAAAAATT CGCCCAAAAT CAAAACGCCT ATGCGAGCAG CGAGAATTTA 180
GACGAGCTTT TAAAGCATGC TAAAATTTCC AGTTTGATGT TTTTAGCTAG GGCGTATTCT 240
AAAGCGGATG TGGAAATGAG CATTGAAATC TTAAAAGGGC TTTTGAATCG CCCCTTAAAA 300
GATGAAGAAA AAATCGCTGT TTTAGATTTA TTGGCTAAAA ATTATTTTAG CGTGGGGTAT 360
TTGCAGAAAA CAAAAGACAC CGTGAAAGAA ATTTTGCGCT TTTCCCCAAG GAATGTGGAA 420
GCGTTGTTGA AGCTTTTGCA TGCGTATGAA TTAGAAAAAG ATTATTCAAA GGCTTTAGAA 480
ACTTTGGAAT GTTTGGAAGA ATTAGAGGTG CCTAAAATTG AAACGATTAA AAATTACCTC 540
TATTTAATGC ATTTAATAGA GAATAAGGAA GATGCGGCTA AAATCTTGCA TGTTTCAAAA 600
GCGTCGTTAG ATTTGAAAAA AATCGCTCTG AATCACTTAA AATCGCATGA TGAAAATCTT 660
TTTTGGCAAG AAATTGATAC AACCGAACGG CTAGAAAATG TGATCGATCT TTTATGGGAT 720
ATGAATATCC CTGCTTTTAT TTTAGAAAAA CATGCCCTTT TGCAGGACAT CGCGCGATCT 780
CAAGGGTTGC TTTTGGATCA CAAACCTTGC CAAATTTTTG AATTAGAGGT TTTACGCGCT 840
CTATTGCATA GCCCTATAAA AGCGAGTCTG ACTTTTGAAT ACCGCTGCAA GCATTGCAAA 900
CAAATCTTTC CTTTTGAAAG CCATAGGTGT CCTGTGTGTT ACCAGTTAGC GTTTATGGAT 960
ATGGTGCTTA AAATCTCTAA AAAAACGCAT GCTATGGGAG TGGATTAA 1008

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1242 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1242

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

ATGAGGAAAA TTTTTTCTTA TATTTCTAAG GTTCTATTAT TTATTGGGGT GGTTTATGCA 60
GAGCCTGATT CTAAAGTGGA AGCCTTAGAA GGGAGGAAGC AAGAGTCTTC TTTGGATAAA 120
AAAATCCGCC AAGAATTGAA GAGTAAGGAA TTGAAGAATA AGGAATTAAA GAATAAGGAT 180
TTGAAAAATA AAGAAGAAAA GAAAGAAACA AAAGCCAAGA GAAAACCCAG AGCAGAAGTC 240
CATCATGGGG ACGCCAAAAA TCCCACTCCA AAGATCACGC CTCCTAAAAT CAAAGGGAGT 300
AGTAAGGGCG TTCAAAATCA AGGCGTTCAA AACAACGCGC CAAAACCTGA AGAAAAAGAT 360
ACAACCCCTC AAGCTACTGA AAAAAATAAG GAAACAAGCC CTAGCTCTCA ATTCAATTCC 420
ATTTTTGGTA ATCCTAATAA CGCTACCAAC AACACCCTTG AAGATAAGGT CGTAGGGGGC 480
ATTTCATTGC TTGTTAATGG TTCGCCTATC ACGCTGTATC AAATCCAAGA AGAGCAAGAA 540
AAATCTAAAG TGAGTAAGGC TCAAGCTAGG GATCGTTTGA TCGCTGAACG CATTAAAAAC 600
CAAGAAATTG AGCGCTTAAA AATCCATGTA GATGATGACA AGCTAGACCA AGAAATGGCG 660
ATGATGGCGC AACAACAAGG CATGGATTTA GACCATTTCA AACAGATGCT TATGGCTGAG 720
GGGCATTATA AACTCTATAG AGATCAACTT AAAGAGCATT TAGAAATGCA AGAATTGTTG 780
CGTAATATTT TGCTCACGAA TGTGGATACC AGCTCTGAAA CCAAAATGCG CGAATATTAC 840
AACAAACACA AGGAGCAATT CAGTATCCCC ACAGAAATAG AAACCGTGCG CTACACTTCC 900

ACCAATCAAG AGGATTTAGA AAGGGCTATG GCAGACCCTA ATTTGGAAGT CCCAGGGGTG 960
AGTAAGGCCA ATGAAAAAAT AGAGATGAAA ACCCTAAACC CTCAAATCGC CCAAGTCTTT 1020
ATTTCGCATG AGCAAGGCTC TTTCACGCCC GTTATGAATG GGGGTGGGGG GCAGTTCATC 1080
ACCTTTTATA TCAAGGAAAA AAGGGGTAAA AATGAAGTGA GCTTCAGTCA GGCCAAGCAA 1140
TTCATCGCCC AAAAATTAGT GGAAGAATCT AAGGATAAGA TTTTAGAAGA GCATTTTGAA 1200
AAATTGCGCG TTAAGTCTAG GATTGTGATG ATCAGAGAGT GA 1242

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 561 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...561

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

ATGATTAAAA GAATTGCTTG TATTTTAAGC TTGAGCGCGA GTTTAGCGTT AGCTGGCGAA 60
GTGAATGGGT TTTTCATGGG TGCGGGTTAT CAACAAGGTC GTTATGGCCC TTATAACAGC 120
AATTACTCTG ATTGGCGTCA TGGCAATGAC CTTTATGGTT TGAATTTCAA ATTAGGTTTT 180
GTAGGCTTTG CCAATAAATG GTTTGGGGCT AGGGTGTATG GCTTTTTAGA TTGGTTTAAC 240
ACTTCAGGGA CTGAACACAC CAAAACCAAT TTGCTCACCT ATGGCGGCGG TGGCGATTTG 300
ATTGTCAATC TCATTCCTTT GGATAAATTC GCTCTAGGTC TCATTGGTGG CGTTCAATTA 360
GCCGGAAACA CTTGGATGTT CCCTTATGAT GTCAATCAAA CCAGATTCCA GTTCTTATGG 420
AATTTAGGCG GAAGAATGCG TGTTGGGGAT CGCAGTGCGT TTGAAGCGGG CGTGAAATTC 480
CCTATGGTTA ATCAGGGTAG CAAAGATGTA GGGCTTATCC GCTACTATTC TTGGTATGTG 540
GATTATGTCT TCACTTTCTA G 561

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 729 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...729

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

ATGAAAAAAT TTTTTTCTCA ATCTTTGTTA GCTCTTATTA TCTCTATGAA TGCGGTATCT 60
GGCATGGATG GTAATGGCGT TTTTTTAGGG GCGGGTTATT TGCAAGGACA GGCGCAAATG 120
CATGCGGATA TTAATTCTCA AAAACAAGCC ACCAACGCTA CGATCAAAGG CTTTGACGCG 180
CTCTTGGGGT ATCAATTTTT CTTTGAAAAA CACTTTGGCT TACGCCTTTA TGGGTTTTTT 240
GACTACGCTC ATGCCAATTC TATTAAGCTT AAAAACCCTA ACTATAATAG CGAAGCGGCG 300
CAAGTGGCTA GTCAAATTCT TGGGAAACAA GAAATCAATC GTTTAACAAA CATTGCCGAT 360
CCCAGAACTT TTGAGCCGAA CATGCTCACT TATGGGGGGG CTATGGACGT GATGGTTAAT 420
GTCATCAATA ACGGCATCAT GAGTTTGGGG GCTTTTGGCG GGATACAATT GGCCGGCAAT 480
TCATGGCTTA TGGCGACACC GAGCTTTGAG GGCATTTTAG TGGAACAAGC CCTTGTGAGC 540
AAGAAAGCCA CTTCTTTCCA ATTTTTATTC AATGTGGGGG CTCGCTTAAG GATCTTAAAA 600
CATTCTAGCA TTGAAGCGGG CGTGAAATTC CCCATGCTAA AGAAAAACCC CTACATCACT 660
GCAAAAAATT TGGATATAGG GTTTAGGCGC GTGTATTCGT GGTATGTGAA TTACGTGTTC 720
ACTTTCTAG 729

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 771 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...771

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

ATGGGATACG CAAGCAAATT AGCTTTAAAG ATTTGTTTGG TAGGTTTATG TTTATTTAGC 60
ACCCTTGGTG CAGAACACCT TGAGCAAAAA GGGAATTATA TTTATAAGGG AGAGGAGGCT 120
TATAATAATA AGGAATATGA GCGAGCGGCT TCTTTTTATA AGAGCGCTAT TAAAAATGGT 180
GAGTCGCTTG CTTATATTCT TTTAGGGATC ATGTATGAAA ATGGTAGGGG TGTACCTAAA 240
GATTACAAGA AAGCGGTTGA ATATTTCCAA AAAGCTGTTG ATAACGATAT ACCTAGAGGG 300
TATAACAATT TGGGCGTGAT GTATAAAGAG GGTAAGGGAG TTCCTAAAGA TGAAAAGAAA 360
GCGGTGGAAT ATTTTAGAAT AGCTACAGAG AAAGGTTATA CTAACGCTTA TATCAACTTA 420
GGCATCATGT ATATGGAGGG CAGGGGAGTT CCAAGTAACT ATGCGAAAGC GACAGAATGT 480
TTTAGAAAAG CGATGCATAA GGGCAATGTG GAAGCTTATA TTCTCCTAGG GGATATTTAT 540

TATAGCGGGA ATGATCAATT GGGTATTGAG CCGGACAAAG ATAAGGCTGT TGTCTATTAT 600
AAAATGGCGG CTGATGTGAG TTCTTCTAGA GCTTATGAAG GGTTGTCAGA GTCTTATCGG 660
TATGGGTTAG GCGTGGAAAA AGATAAAAAA AAGGCTGAAG AATACATGCA AAAAGCATGC 720
GATTTTGACA TTGATAAAAA TTGTAAGAAA AAGAACACTT CAAGCCGATA A 771

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1974 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1974

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

ATGAGAAAAC TATTCATCCC ACTTTTATTA TTCAGCGCTT TAGAAGCGAA CGAGAAAAAC 60
GGCTTTTTCA TAGAAGCCGG CTTTGAAACT GGGCTATTAG AAGGCACACA AACGCAAGAA 120
AAAAGACACA CCACCACAAA AAACACTTAC GCAACTTACA ATTATTTACC CACAGACACG 180
ATTTTAAAAA GAGCGGCTAA TTTATTCACC AATGCCGAAG CGATTTCAAA ATTAAAATTC 240
TCATCTTTAT CCCCTGTTAG AGTGTTGTAT ATGTATAATG GTCAATTAAC TATAGAAAAC 300
TTCTTGCCTT ATAATTTAAA TAATGTTAAG CTTAGTTTTA CAGACGCTCA AGGCAATGTG 360
ATCGATCTAG GCGTGATAGA GACTATCCCC AAACACTCTA AGATTGTTTT GCCCGGAGAG 420
GCATTTGATA GTCTAAAAAT TGACCCCTAT ACTTTATTTC TTCCAAAAAT TGAAGCCACT 480
AGCACTTCTA TTTCTGACGC TAACACGCAG AGGGTGTTTG AAACGCTCAA TAAGATTAAG 540
ACAAATTTGG TCGTAAATTA TAGGAATGAA AACAAATTTA AAGATCACGA AAATCATTGG 600
GAAGCCTTTA CCCCACAAAC CGCAGAAGAA TTCACTAATT TAATGTTGAA CATGATCGCT 660
GTTTTAGACT CCCAATCTTG GGGCGATGCG ATCTTAAACG CTCCTTTTGA GTTCACTAAC 720
AGCCCAACAG ATTGCGATAA TGATCCTTCA AAATGCGTAA ATCCTGGGAC AAACGGGCTT 780
GTCAATTCTA AAGTCGATCA AAAATATGTG TTAAACAAAC AAGACATTGT CAATAAATTT 840
AAAAACAAAG CGGATCTTGA TGTAATTGTT TTAAAGGATT CAGGGGTTGT AGGGCTTGGG 900
AGTGATATTA CCCCTAGCAA CAATGATGAT GGCAAGCATT ATGGCCAGTT AGGGGTAGTA 960
GCTTCTGCTT TAGATCCTAA AAAACTCTTT GGCGATAACC TTAAGACTAT CAATTTAGAG 1020
GATTTAAGAA CCATCTTGCA TGAATTCAGC CACACTAAAG GCTATGGGCA TAACGGGAAT 1080
ATGACCTATC AAAGAGTGCC GGTAACGAAA GATGGTCAAG TGGAAAAGGA TAGTAATGGC 1140
AAGCCAAAAG ATTCTGATGG CCTCCCCTAT AATGTGTGTT CGCTTTATGG GGGATCCAAT 1200
CAGCCCGCTT TCCCTAGCAA CTACCCTAAT TCCATCTATC ACAATTGTGC GGATGTCCCG 1260
GCTGGCTTTT TAGGGGTAAC AGCAGCGGTT TGGCAGCAGC TCATCAATCA AAACGCCTTG 1320
CCGATCAACT ACGCTAACTT GGGGAGTCAA ACAAACTACA ACCTAAACGC TAGTTTAAAC 1380
ACGCAAGATT TAGCCAATTC CATGCTCAGC ACCATCCAAA AAACCTTTGT AACTTCTAGC 1440
GTTACCAACC ACCATTTTTC AAACGCATCG CAAAGTTTTA GAAGCCCTAT TTTAGGGGTT 1500
AACGCTAAAA TAGGCTATCA AAACTACTTT AATGATTTCA TAGGGTTGGC TTATTATGGC 1560
ATCATCAAAT ACAATTACGC TAAAGCTGTT AATCAAAAAG TCCAGCAATT GAGCTATGGT 1620

GGGGGGATAG ATTTGTTATT GGATTTCATC ACCACTTACT CCAATAAAAA TAGCCCTACA 1680
GGCATTCAAA CCAAAAGGAA TTTTTCTTCA TCTTTTGGTA TCTTTGGGGG GTTAAGGGGC 1740
TTGTATAACA GCTATTATGT GTTGAACAAA GTCAAAGGAA GCGGCAATTT AGATGTGGCT 1800
ACCGGGTTGA ACTACCGCTA TAAGCATTCT AAATATTCTG TAGGGATTAG CATCCCTTTA 1860
ATCCAAAGAA AAGCTAGCGT CGTTTCTAGC GGTGGCGATT ATACGAACTC TTTTGTTTTC 1920
AATGAAGGGG CTAGCCACTT TAAGGTGTTT TTCAATTACG GGTGGGTGTT TTAG 1974

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 504 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...504

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

ATGAAATTGG TGAGTCTTAT TGTAGCGTTA GTTTTTTGTT GTTTTTTAGG GGCTGTAGAG 60
TTGCCTGGAG TTTATCAAAC TCAAGAATTT TTATACATGA AAAGCTCTTT TGTGGAGTTT 120
TTTGAGCATA ACGGGAAGTT CTATGCCTAT GGTATTTCTG ATGTGGATGG CTCTAAAGCC 180
AAAAAAGACA AACTCAATCC TAACCCAAAG CTAAGGAATC GCAGCGATAA AGGCGTGGTG 240
TTTTTAAGCG ATTTGATTAA GGTTGGGGAA CAATCTTATA AAGGCGGTAA GGCGTATAAT 300
TTTTATGACG GCAAGACCTA CCATGTGAGA GTCACTCAAA ATTCAAACGG GGATTTGGAA 360
TTCACTTCAA GCTATGACAA ATGGGGGTAT GTGGGCAAAA CCTTCACCTG GAAACGCCTG 420
AGCGATGAAG AAATCAAAAA TCTAAAGCTC AAGCGTTTTA ACTTGGACGA AGTCCTTAAA 480
ACCCTCAAAG ATAGCCCTAT TTAA 504

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 885 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...885

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

ATGAGTAATC AAGCGAGCCA TTTGGATAAT TTTATGAACG CTAAAAATCC CAAAAGTTTT 60
TTTGATAATA AGGGGAATAC CAAATTCATC GCTATCACAA GCGGTAAGGG GGGCGTGGGG 120
AAATCCAACA TTAGCGCTAA TTTAGCTTAC TCTTTATACA AGAAAGGTTA TAAGGTAGGG 180
GTATTTGATG CGGATATTGG TTTAGCGAAT TTAGATGTCA TTTTTGGGGT GAAAACCCAT 240
AAAAATATCT TGCATGCCTT AAAAGGCGAA GCCAAATTGC AAGAAATCAT TTGCGAGATT 300
GAACCCGGGC TTTGCTTAAT CCCTGGGGAT AGCGGCGAAG AAATTTTAAA ATACATCAGC 360
GGCGCGGAAG CTTTGGATCG ATTCGTAGAT GAAGAGGGGG TTTTAAGCTC TTTAGATTAT 420
ATTGTGATTG ATACGGGTGC TGGGATTGGG GCCACTACGC AAGCGTTTTT GAATGCGAGC 480
GATTGCGTGG TGATTGTTAC CACACCCGAT CCTTCAGCGA TTACCGATGC GTATGCATGC 540
ATTAAAATCA ACTCCAAGAA TAAAGATGAA TTGTTCCTTA TCGCTAACAT GGTAGCCCAA 600
CCTAAAGAAG GCAGGGCGAC TTATGAAAGG CTATTCAAGG TGGCTAAAAA CAATATCGCT 660
TCATTAGAAT TGCACTATTT AGGGGCGATT GAAAACAGCT CCTTATTGAA ACGCTATGTG 720
AGGGAGCGAA AGATTTTGAG GAAAATAGCC CCTAACGATT TGTTTTCGCA ATCCATTGAC 780
CAGATAGCGA GCCTTTTAGT TTCTAAACTA GAAACCGGCA CTTTAGAAAT ACCAAAAGAA 840
GGTTTAAAAA GCTTTTTTAA AAGGCTTTTG AAGTATTTGG GGTAG 885

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1119 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1119

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TTGGAACCTT CAAGAAATCG CCTAAAACAT GCCGCCTTTT TTGTGGGGCT TTTTATCGTT 60
TTGTTTTTAA TTATAATGAA GCACCAAACC TCCCCCTATG CTTTCACGCA TAATCAAGCC 120
CTTGTCACTC AAACCCCCCC CTATTTCACG CAACTCACTA TCCCTAAACC AAATGACGCT 180
TTAAGCGCGC ATGCGAGCTC TTTAATCAGC TTGCCTAACG ACAATCTTTT GAGCGCTTAT 240
TTTAGCGGCA CTAAAGAAGG GGCAAGGGAT GTGAAAATCA GCGCGAATCT TTTTGACAGC 300
AAGACTAATC GCTGGAGCGA AGCCTTCATT CTTTTAACCA AAGAAGAGCT TTCTCATCAT 360
TCGCATGAAT ACATCAAAAA ATTAGGTAAC CCCTTGCTTT TTTTGCATGA TAATAAAATT 420
TTGTTGTTTG TCGTAGGGGT GAGCATGGGC GGGTGGGCCA CTTCTAAAAT CTATCAATTT 480

GAAAGCGCTT TAGAGCCGAT TCATTTTAAG TTTGCGCGAA AACTCTCTTT AAGCCCTTTT 540
TTAAATTTGA GCCATTTAGT AAGGAATAAG CCTTTAAACA CCACTGATGG CGGGTTTATG 600
CTACCACTCT ATCACGAATT AGCCACCCAA TACCCCTTGT TGTTGAAATT TGACCAACAA 660
AATAACCCAA GAGAGCTTTT AAGGCCTAAT ACCTTAAACC ACCAGCTCCA ACCAAGCTTA 720
ACCCCCTTTA AAGACTGCGC TGTCATGGCG TTTAGAAACC ATTCTTTTAA AGATAGCCTC 780
ATGCTAGAAA CCTGTAAAAC CCCCACTGAT TGGCAAAAAC CCATTTCTAC AAATCTTAAA 840
AACTTAGATG ATTCTTTAAA TTTACTCAAT TTAAATGGAA TATTGTATTT GATCCACAAC 900
CCTAGCGATT TATCACTGCG TCGTAAAGAA CTTTGGCTTT CTAAATTAGA AAACTCCAAC 960
TCGTTTAAAA CCTTAAAAGT TTTGGATAAA GCGAATGAAG TGAGTTACCC AAGCTATAGC 1020
CTTAATCCGC ATTTTATAGA TATTGTCTAT ACTTACAACC GCTCTCATAT CAAACACATC 1080
CGTTTCAATA TGGCTTATTT AAATTCCCTT CTCAAGTGA 1119

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2937 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...2937

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

ATGAAGAAAA GAAAACATGT ATCCAAGAAA GTGTTTAATG TCATTATCTT GTTTGTGGCA 60
GTATTCACTC TTTTAGTCGT CATTCACAAA ACCCTTTCAA ACGGCATTCA CATACAAAAT 120
TTAAAAATTG GAAAACTTGG CATTTCTGAA TTATACTTAA AACTCAATAA CAAGCTTTCT 180
TTGGAAGTTG AGCGGGTTGA TCTCTCTTCT TTCTTCCATC AAAAACCCAC TAAAAAGCGT 240
TTAGAAGTTT CTGATTTGAT TAAAAATATC CGTTATGGCA TTTGGGCGGT GTCTTATTTT 300
GAAAAACTTA AAGTCAAAGA AATCATTTTA GACGATAAAA ATAAAGCCAA TATCTTTTTT 360
GATGGGAATA AATACGAGTT AGAATTTCCA GGAATCAAAG GGGAATTTTC CCTAGAAGAC 420
GATAAAAATA TCAAGCTTAA AATCATCAAT TTGCTTTTTA AAGATGTTAA AGTCCAAGTG 480
GATGGCAACG CCCACTATTC ACCCAAAGCC AGGAAAATGG CGTTCAATTT GATTGTCAAG 540
CCCTTAGTTG AACCCAGCGC TGCAATTTAT TTGCAAGGGC TAACCGATTT AAAAACCATA 600
GAATTAAAAA TTAACACTTC TCCAATGAAA AGCCTAGCGT TTTTAAAGCC TCTTTTCCAA 660
CGCCAATCGC AAAAAAATTT AAAAACGTGG ATTTTTGACA AGATCCAATT TGCCAGCTTT 720
AAGATTGATA ACGCTTTAAT CAAGGCTAAT TTCACTCCTA GCGAGTTTAT CCCATCGCTT 780
TTGGAAAATT CTGTAGTTAA AGCCACTTTG ATTAAGCCTT CAGTCGTTTT TAATGATGGC 840
TTATCGCCCA TTAAAATGGA TAAAACCGAA TTGATTTTCA AAAACAAACA GCTCCTCATA 900
CAGCCCCAAA AAATCACTTA TGAAACCATG GAATTAACCG GCTCTTACGC CACTTTTTCC 960
AATTTGTTAG AAGCCCCTAA GTTGGAGGTT TTTTTAAAAA CGACCCCTAA TTATTATGGC 1020
GATAGCATTA AGGATTTATT GAGCGCTTAT AAAGTCGTTT TACCTTTGGA TAAAATCAGC 1080
ATGCCATCTA GCGCGGATTT GAAGCTCACT TTGCAATTCT TAAAAAACAC CGCCCCCTTA 1140
TTTAGCGTTC AAGGCAGCGT TAATTTGCAA GAAGGCACTT TCTCGCTCTA TAATATCCCC 1200

CTTTACACGC AAAGCGCTCA AATCAATTTG GACATCGCCC AAGAATACCA ATACATCTAC 1260
ATAGACACGA TCCACACGCG CTATGCAAAC ATGCTGGATT TAGACGCTAA AATCGCTTTA 1320
GATTTAGGTC AAAAAAACCT TTCTTTGGAT TCTTTAGTCC ATAAAATCCA AGTCAATACC 1380
AATAACAATA TCAACATGCG CTCTTATGAT CCCAATAACA CTCAAGAAGA TCCGCAAACT 1440
AACTTTACTT TGGATCTAAA AAGCTTGCAT TCTATCATTC AAGAGGGTGA AAATTCAGAA 1500
GTTTTTAGAA GAAAAATCAT AGACACCATT AAAGCCCAAA GCGAAGATAA ATTCACTAAA 1560
GATGTTTTTT ACGCCACAGG AGACACTCTC AAAAGCCTGT CGTTGAGTTT TGATTTTTCC 1620
AACCCCGATC ACATACAATG GAGCGTGCCA CAACTCTTAT TAGAAGGCGA ATTTAAAGAT 1680
AACGCCTATA CTTTTAAGAT CAAAGATTTG AAAAAGATCA AGCCCTATTC CCCCATTATG 1740
GACTATATTG CCCTAAAAGA CGGCTCTTTA GAGGTTTCTA CGAGCGATTT TGTCAATATT 1800
GATTTTTTTG CTAAAGATTT GAAAATCAAC CTCCCCATTT ATAGGAGCGA TGGATCGCAT 1860
TTTGATTCTT TTTCTTTATT TGGCTCTATC AATAAAGATG AAATTTCTGT CTATACTCCA 1920
AGCAAAAGCA TATCCATAAA AGTTAAGGGG GATCAAAAGG ATATTACCCT TAATAACATT 1980
GATTTGAGTA TTGATGATTT CTTGGATAGT AAAATGCCAG CTATTGCGGG ATTATTCTCA 2040
AAAGAACGAA AAGAAAAGCC TAGCTCTAAA GAAATCCAAG ATGAAGATGT TTTCATTAGC 2100
GCCAAACAAC GCTATGAAAA AGCCCACAAA ATTATCCCCA TCTCTACACG CATCCATGCT 2160
AAAGATGTCG TGCTGATCTA TAAAAAAATG CCTTTTCCTT TAGAAAATCT TGATATTGTC 2220
GCTCAAGACG ATAGGGTGAA AATTGATGGC AATTATAAAA ACGCCATGAT CATGGCGGAT 2280
TTAGTGCATG GGGCTTTGTA TCTTAAGGCT CATAATTTTA GCGGGGATTA TATCAACACC 2340
ATTCTTCAAA AAGATTTCGT AGAAGGAGGC TTATTCACGC TTATTGGGGC TCTTGAAGAT 2400
CAGGTTTTCA ATGGCGAATT GAAATTCCAA AACACAAGCT TAAAGAATTT CGCCCTCATG 2460
CAAAACATGG TCAATCTCAT CAACACCATT CCCTCCCTCA TTGTCTTTAG AAACCCTCAT 2520
TTAGGGGCTA ATGGCTATCA AATCAAAACC GGCTCCGTTG TGTTTGGGAT CACTAAAGAA 2580
TATTTAGGGT TAGAAAAAAT TGATCTTGTC GGCAAAACGC TTGATATTGC TGGCAATGGA 2640
ATCATTGAAT TAGACAAAAA CAAATTAGAT TTAAACTTAG AAGTTTCCAC TATCAAGGCT 2700
TTGAGTAATG TCTTAAATAA AATCCCTATC GTGGGCTATC TCGTTTTAGG AAAAGGAGGT 2760
AAAATCACCA CTAACGTGAA TGTCAAAGGC ACGTTGGATA AGCCTAAAAC CCAAGTAACT 2820
TTAGCGTCAG ATATTATCCA AGCGCCTTTT AAAATCTTAC GCCGTATTTT CACGCCTATT 2880
GACATCATCG TGGATGAAGT CAAGAAAAAC ATTGATTCAA AAAGGAAATT AAAATGA 2937

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1434 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1434

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

ATGAATACTA TTATAAGATA TGCGAGTTTA TGGGGCTTGT GTATTACTCT AACTCTAGCG 60
CAAACCCCCT CTAAAACCCC TGATGAAATC AAGCAAATCC TTAACAATTA TAGCCATAAG 120

AATTTAAAGC TCATTGATCC GCCGACAAGT TCTTTAGAAG CGACACCGGG TTTTTTACCC 180
TCGCCTAAAG AAACAGCGAC CACGATCAAT CAAGAGATCG CTAAATACCA TGAAAAAAGC 240
GATAAAGCCG CTTTGGGGCT TTATGAATTG CTAAAGGGGG CTACCACCAA TCTCAGTTTG 300
CAAGCGCAAG AACTCAGTGT CAAGCAAGCG ATGAAGAACC ACACCATCGC CAAAGCGATG 360
TTTTTGCCTA CTTTGAACGC GAGTTATAAT TTTAAAAATG AAGCTAGGGA TACTCCAGAA 420
TATAAGCATT ATAACACCCA ACAACTCCAA GCTCAAGTCA CATTGAATGT GTTTAATGGC 480
TTTAGCAATG TGAATAATGT CAAAGAAAAG TCTGCGACTT ACCGATCCAC TGTGGCTAAT 540
TTAGAATATA GCCGCCAAAG CGTGTATTTG CAAGTGGTGC AACAATACTA CGAGTATTTT 600
AACAATCTCG CTCGCATGAT CGCTTTGCAA AAGAAATTAG AGCAAATCCA AACGGACATT 660
AAAAGGGTTA CTAAGCTCTA TGACAAAGGG CTGACCACGA TTGATGATTT ACAAAGCTTA 720
AAAGCGCAAG GGAATTTGAG CGAATACGAT ATTTTGGACA TGCAATTTGC TTTGGAGCAA 780
AACCGCTTGA CTTTAGAATA CCTCACTAAC CTCAGTGTGA AAAATTTGAA AAAGACCACG 840
ATTGATGCGC CTAATTTGCA ATTAAGAGAA AGGCAGGATT TGGTTTCTTT AAGGGAGCAG 900
ATTTCTGCAC TCAGATACCA AAACAAGCAA CTCAATTATT ACCCCAAGAT AGATGTGTTT 960
GACTCATGGC TTTTTTGGAT CCAAAAACCC GCTTATGCCA CAGGGCGTTT TGGGAATTTC 1020
TACCCAGGTC AGCAAAATAC GGCTGGGGTT ACTGCGACTT TGAATATTTT TGATGATATA 1080
GGGTTGAGCT TGCAAAAACA ATCCATCATG CTAGGCCAAT TAGCGAATGA AAAGAATTTA 1140
GCGTATAAAA AATTGGAGCA AGAAAAAGAC GAACAGCTTT ACAGAAAGTC GCTTGATATT 1200
GCCAGAGCTA AGATTGAATC TTCAAAGGCT AGTTTGGATG CGGCCAATCT TTCTTTTGCC 1260
AATATTAAAA GGAAATACGA CGCTAATTTA GTGGATTTCA CTACCTATTT AAGGGGCTTA 1320
ACCACGCGCT TTGATGCAGA AGTGGCTTAC AATTTAGCGC TCAACAATTA CGAAGTGCAA 1380
AAAGCCAATT ACATTTTTAA CAGCGGGCAT AAAATAGACG ACTATGTGCA TTAA 1434

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1239 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1239

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

ATGCTATCTT TTATAAGCGC GTTTGATAAA AGGGGCGTTT CAATACGCCT TCTAACAGCC 60
TTGTTACTGC TTTTTAGTTT GGGTTTGGCT AAAGATTTAG AAATCCAAAC TTTTGTGGCT 120
AAATACCTTT CTAAAAATCA AAAAATACAA GCCCTACAGG AGCAAATTGA CGCTTTAGAT 180
TCTCAAGAAA AAGTCGTTAG CAAATGGGAT AACCCTATTT TGTATTTAGG CTATAACAAC 240
GCTAACGTGA GCGATTTTTT CAGGCTGGAT AGCACCTTAA TGCAAAACAT GAGCTTGGGT 300
TTGTCTCAAA AAGTGGATTT AAATGGTAAA AAACTCACGC AGTCTAAAAT GATCAATTTA 360
GAAAAACAAA AAAAAATATT AGAGCTTAAA AAAACCAAGC AGCAATTGGT GATTAATTTA 420
ATGATAAACG GCATTGAAAA CTATAAAAAC CAACAAGAAA TAGAGCTTTT AAACACAGCG 480
ATTAAAAATT TAGAAAACAC CCTCTATCAA GCCAACCATT CCAGTTCGCC CGATTTAATA 540

GCGATCGCCA AGTTAGAAAT TTTAAAATCG CTATTAGAAA TCCAAAAAAA CGATTTAGAA 600
GTAGCGCTCT CTAGCAGCCA TTATTCCATG GGCGAATTGA CTTTTAAAGA AAACGAGATT 660
TTAAGCATTG CCCCTAAAAA TTTTGAATTC AATAACGAGC AAGAGCTGCA TAACATTAGC 720
GCCACTAATT ACGATATTGC GATCGCCAGG CTTGATGAAG AAAAAGCACA AAAAGACATC 780
ACTCTGGCTA AAAAAAGCTT TTTAGAAGAC ATAAACGTTA CCGGGGGTGTA TTATTTCCGC 840
TCCAAACAAT ACTATAACTA CGACATGTTT AGCGTCGCTT TGTCTATCCC TTTACCTCTT 900
TATGGCAAGC AGGCTAAATT AGTGGAGCAA AAGAAAAAAG AAAGCTTGGC GTTTAAAAGC 960
GAAGTGGAAA ACGCCAAAAA CAAAACGCGC CACCTGGCCC TAAAACTCCT TAAAAAATTA 1020
GAAACCTTGC AAAAAAACCT GGAATCGATC AATAAAATCA TCAAACAGAA TGAAAAAATC 1080
GCGCAAATTT ATGCGCTTGA TTTGAAAACT AATGGCGATT ACAACGCTTA TTACAACGCC 1140
TTGAATGACA AAATCACTAT TCAAATCACC CAGCTTGAAA CCTTAAGCGC TCTAAATAGT 1200
GCTTATTTGT CCTTACAAAA TCTCAAAGGA TTAGAATGA 1239

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 414 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...414

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

ATGCGTATAG TTAGAAATTT ATTTCTTGTA TCGTTTGTGG CGTATAGTAG TGCGTTCGCA 60
GCGGATTTAG AAACCGGAAC CAAAAACGAC AAAAAGAGCG GTAAAAAATT TTACAAACTC 120
CATAAAAACC ATGGCTCAGA AACCGAGACT AAAAACGATA AAAAGCTTTA TGATTTCACT 180
AAAAATAGCG GATTAGAAGG CGTGGATTTA GAAAAAAGCC CTAACCTTAA AAGCCATAAA 240
AAAAGCGATA AAAAGTTTTA TAAACAACTC GCTAAAAACA ATATCGCTGA AGGGGTGAGC 300
ATGCCGATTG TGAATTTCAA TAAAGCCCTA TCTTTTGGGC CTTATTTTGA AAGGACTAAA 360
AGCAAAAAAA CCCAATACAT GGACGGCGGG TTGATGATGC ACATCCGTTT TTAA 414

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 930 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...930

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TTGATGCCAC AAAACCAGCT TGTGATCACC ATCATTGATG AATCAGGCTC TAAGCAACTC 60
AAATTTTCTA AAAATTTAAA ACGCAACCTC ATCATTTCTG TTGTCATTCT TTTATTGATC 120
GTGGGGCTTG GCGTGGGGTT TTTAAAATTT TTAATCGCTA AAATGGATAC GATGACAAGC 180
GAGAGGAATG CGGTTTTAAG GGATTTTAGG GGTTTGTATC AAAAAAATTA CGCCCTAGCG 240
AAAGAGATTA AAAACAAGCG AGAAGAGCTT TTTATTGTGG GGCAAAAGAT CCGTGGGCTA 300
GAATCCTTGA TTGAAATCAA AAAGGGGGCT AATGGGGGAG GGCATCTCTA TGATGAAGTG 360
GATTTAGAAA ATTTGAGCTT AAATCAAAAA CATTTAGCAC TCATGCTCAT TCCTAATGGC 420
ATGCCCCTAA AAACTTATAG CGCTATCAAA CCCACTAAAG AAAGGAACCA CCCCATTAAA 480
AAGATTAAGG GCGTTGAATC CGGGATCGAT TTTATCGCGC CATTGAACAC GCCTGTGTAT 540
GCGAGCGCTG ATGGGATTGT GGATTTTGTG AAGACTCGTT CTAATGCGGG GTATGGGAAC 600
TTGGTGCGCA TTGAACATGC GTTTGGTTTC AGCTCCATTT ATACGCACTT AGATCATGTC 660
AATGTGCAGC CTAAAAGCTT CATCCAAAAA GGGCAGTTGA TTGGCTATAG CGGGAAGAGC 720
GGTAATAGCG GCGGCGAAAA ATTGCATTAT GAAGTGCGGT TTTTGGGTAA AATTTTAGAC 780
GCAGAAAAAT TCCTAGCATG GGATTTGGAT CATTTTCAAA GCGCTTTAGA AGAAAATAAA 840
TTTATTGAAT GGAAGAATCT GTTTTGGGTT TTAGAAGACA TCGTCCAGCT CCAAGAGCAT 900
GTGGATAAAG ACACCTTAAA AGGTCAGTAG 930

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 999 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...999

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GTGCTATATT TTTTAACCAG TTTATTTATT TGCTCTTTGA TTGTTTTGTG GTCTAAAAAA 60
TCCATGCTCT TTGTGGATAA CGCTAATAAA ATCCAAGGCT TCCATCATGC AAGAACCCCA 120
CGAGCCGGGG GGCTTGGGAT CTTTCTTTCT TTTGCGTTGG CTTGTTATCT TGAACCTTTT 180

GAGATGCCTT TTAAGGGGCC TTTTGTTTTC TTAGGGCTAT CGCTAGTGTT TTTGAGCGGT 240
TTTTTAGAAG ACATTAACCT TTCATTAAGC CCCAAAATAC GCCTTATTTT GCAAGCTGTA 300
GGGGTCGTTT GCATCATTTC ATCAACGCCT TTAGTGGTGA GCGATTTTTC GCCCCTTTTT 360
AGCTTGCCTT ATTTCATCGC TTTTTTATTC GCTATTTTTA TGCTGGTGGG TATCAGTAAC 420
GCTATTAATA TCATTGACGG GTTTAACGGG CTTGCATCTG GGATTTGCGC GATCGCGCTT 480
TTAGTCATTC ATTATATAGA CCCTAGCAGT TTGTCTTGTT TGCTCGCTTA CATGGTGCTT 540
GGGTTTATGG TGTTAAATTT CCCTTCAGGA AAGATTTTTT TAGGCGATGG GGGGGCGTAT 600
TTTTTGGGTT TGGTGTGCGG GATTTCTCTC TTGCATTTGA GTTTGGAGCA AAAAATCAGC 660
GTGTTTTTTG GGCTCAATTT AATGCTTTAT CCGGTCATAG AGGTGCTTTT TAGTATCCTT 720
AGGCGCAAAA TAAAACGCCA GAAAGCCACC ATGCCGGATA ATTTGCATTT GCACACCCTT 780
TTATTTAAAT TCTTGCAACA ACGCTCTTTC AATTACCCTA ACCCTTTATG CGCGTTTATC 840
CTTATTCTAT GCAACCTGCC TTTTATTTTA ATAAGCGTTT TGTTTCGCTT GGACGCTTAT 900
GCGCTCATTG TGATTAGCCT AGTCTTTATC GCATGCTATT TAATAGGCTA TGCTTATTTG 960
AATAGGCAAG TTTGCGCTTT AGAAAAGCGG GCGTTTTAA 999

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 816 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...816

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

ATGAACATAT TCAAGCGTAT TATTTGCGTA ACCGCTATTG TTTTAGGTTT TTTTAACCTT 60
TTAGACGCCA AACACCACAA AGAAAAAAAA GAAGACCACA AAATCACTCG TGAGCTTAAA 120
GTGGGCGCTA ACCCTGTGCC GCATGCGCAA ATCTTGCAAT CAGTTGTGGA TGATTTGAAA 180
GAGAAAGGGA TCAAATTAGT GATCGTGTCT TTTACGGATT ATGTGTTGCC TAATTTAGCG 240
CTCAATGACG GCTCTTTAGA CGCGAATTAC TTCCAGCACC GCCCTTATTT GGATCGGTTT 300
AATTTGGACA GAAAAATGCA CCTTGTTGGT TTGGCCAATA TCCATGTGGA GCCTTTAAGA 360
TTTTATTCTC AAAAAATCAC AGACATTAAA AACCTTAAAA AAGGCTCAGT GATTGCTGTG 420
CCAAATGATC CGGCCAATCA AGGCAGGGCG TTGATTTTAC TCCATAAACA AGGCCTTATC 480
GCTCTCAAAG ACCCAAGCAA TCTATACGCT ACGGAGTTTG ATATTGTCAA AAATCCTTAC 540
AACATCAAAA TCAAACCCCT AGAAGCTGCG TTATTGCCTA AGGTTTTAGG GGATGTGGAT 600
GGGGCTATCA TAACAGGGAA TTATGCCTTG CAAGCAAAAC TCACCGGAGC CTTATTTTCA 660
GAAGATAAGG ACTCGCCTTA TGCTAATCTT GTAGCCTCTC GTGAGGATAA TGCGCAAGAT 720
GAAGCGATAA AAGCGTTGAT TGAAGCCTTA CAGAGCGAAA AGACCAGGAA ATTCATTTTG 780
GATACCTATA AGGGGGCGAT TATCCCGGCT TTTTAA 816

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 951 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...951

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

ATGCAAGAAT TCAGTTTGTG GTGCGATTTT ATAGAAAGGG ATTTTTTAGA AAACGATTTT 60
TTAAAGCTCA TCAATAAGGG GGCTATTTGC GGGGCGACGA GTAACCCTAG TTTGTTTTGC 120
GAAGCGATCA CAAAAAGCGC GTTTTATCAA GATGAAATCG CTAAACTCAA AGGCAAAAAA 180
GCTAAAGAAA TTTATGAAAC TCTGGCACTA AAGGATATTT TACAAGCCTC TAGCGCGTTA 240
ATGCCTTTGT ATGAAAAAGA CCCTAACAAC GGCTACATCA GCCTAGAAAT TGACCCCTTT 300
TTAGAAGACG ATGCGATTAA AAGCATTGAT GAAGCCAAGC GGTTATTCAA AACATTAAAC 360
CGCCCCAATG TGATGATTAA AGTCCCGGCG AGTGAAAGCG CTTTTGAAGT CATTAGCGCT 420
CTGGCTCAAG CCTCTATCCC CATTAATGTA ACTTTAGTCT TTTCGCCTAA AATTGCCGGT 480
GAAATCGCTC AAATCTTAGC CAAAGAAGCA CGAAAAAGAG CGGTCATTAG CGTGTTTGTC 540
TCACGATTTG ACAAAGAAAT AGACCCACTA GTGCCACAAA ATTTGCAAGC TCAAAGTGGG 600
ATCATGAACG CTACCGAGTG TTATTATCAA ATCAACCAGC ATGCTAATAA GCTAATAAGC 660
ACCCTTTTTG CATCCACCGG CGTTAAATCT AATTCTTTAG CTAAAGATTA CTACATTAAA 720
GCGCTGTGTT TTAAAAACTC TATCAACACA GCCCCCCTAG ACGCCCTAAA CGCTTATTTG 780
CTTGACCCAA ACACCGAGTG TCAAACCCCT TTAAAAATCA CAGAAATTGA AGCGTTCAAA 840
AAAGAATTAA AAACGCACAA TATTGATTTA GAAAACACCG CCCAAAAACT CCTTAAAGAA 900
GGCTTGATAG CGTTCAAACA ATCCTTTGAA AAGCTTTTAA GCAGTTTTTG A 951

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 783 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...783

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

ATGAAAACAA ATGGTCATTT TAAGGATTTT GCATGGAAAA AATGCTTTTT AGGCGCGAGC 60
GTGGTGGCTT TATTAGTGGG GTGTAGCCCG CATATTATTG AAACCAATGA AGTTGCTTTG 120
AAATTGAATT ACCATCCAGC TAGCGAGAAA GTTCAAGCGT TAGATGAAAA GATTTTACTT 180
TTAAGGCCAG CTTTCCAATA CAGCGATAAT ATTGCTAAAG AGTATGAAAA CAAATTCAAG 240
AATCAAACCA CGCTTAAAGT TGAAGAGATC TTGCAAAATC AGGGCTATAA GGTTATTAAT 300
GTGGATAGCA GCGATAAAGA CGATTTTTCT TTTGCGCAAA AAAAAGAAGG GTATTTGGCT 360
GTCGCTATGA ATGGCGAAAT TGTTTTACGC CCCGATCCTA AAAGGACCAT ACAGAAAAAA 420
TCAGAACCCG GGTTATTATT CTCCACTGGT TTGGATAAAA TGGAAAGGGT TTTAATCCCG 480
GCTGGGTTTG TCAAGGTTAC CATACTAGAG CCTATGAGTG GGGAATCTTT GGATTCTTTT 540
ACGATGGATT TGAGCGAGTT GGACATCCAA GAAAAATTCT TAAAAACCAC CCATTCAAGC 600
CATAGCGGAG GGTTAGTTAG CACTATGGTT AAGGGGACGG ATAATTCTAA TGACGCAATT 660
AAGAGCGCTT TGAATAAGAT TTTTGCAAGT ATCATGCAAG AAATGGATAA GAAACTCACT 720
CAAAGGAATT TAGAATCTTA TCAAAAAGAC GCCAAGGAAT TAAAAAACAA GAGAAACCGA 780
TAA 783

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4149 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...4149

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TTGAATTTTA ATAACCTTAC GGCTAATGGG GCGTTAAATT TTAATGGTTA TGCGCCCTCT 60
TTAACTAAGG CTTTAATGAA TGTCAGCGGG CAGTTTGTTT TAGGGAATAA TGGGGATATT 120
AATTTATCTG ACATCAATAT CTTTGACAAC ATCACAAAAT CTGTAACTTA CAACATCTTA 180
AACGCTCAAA AAGGGATTAC TGGCATTAGT GGGGCTAATG GCTATGAAAA AATCCTTTTT 240
TATGGCATGA AAATCCAAAA CGCTACCTAT AGCGATAATA ACAACATCCA AACTTGGTCG 300
TTTATAAACC CTCTCAATTC TTCTCAAATC ATTCAAGAGA GCATTAAAAA TGGGGATCTA 360
ACCATAGAAG TTTTAAATAA CCCTAACTCG GCTTCCAACA CTATTTTTAA TATCGCTCCT 420
GAGCTTTATA ATTACCAAGA TTCTAAGCAA AATCCTACCG GCTATAGCTA TGATTATAGC 480
GACAATCAAG CAGGCACTTA TTACTTGACA AGCAACATTA AAGGTCTTTT CACCCCTAAA 540
GGCTCTCAAA CGCCTCAAAC CCCAGGCACT TATAGCCCAT TTAACCAGCC TTTGAATAGT 600
TTGAATATCT ACAATAAGGG TTTTTCTAGC GAGAATTTAA AAACGCTTTT AGGGATCCTT 660

TCTCAAAATT CCGCCACCTT AAAAGAAATG ATTGAATCCA ACCAACTAGA CAATATCACT 720
AACATTAATG AAGTGTTGCA ACTCTTAGAT AAGATTAAAA TCACCCAAGC GCAAAAGCAA 780
GCGCTCCTAG AAACGATCAA CCATTTGACT GACAACATCA ATCAAACCTT TAATAACGGG 840
AATCTCGTTA TAGGCGCTAC CCAAGATAAT GTTACAAACT CTACTAGCTC TATATGGTTT 900
GGGGGCAATG GCTATAGCAG CCCTTGCGCG CTAGATAGCG CCACTTGTTC TTCTTTTAGA 960
AACACTTACT TGGGGCAATT ATTAGGCTCA ACTTCCCCTT ATTTAGGCTA CATTAACGCT 1020
GATTTTAAAG CTAAAAGCAT TTATATTACC GGGACAATTG GAAGTAGTAA CGCTTTTGAA 1080
AGCGGAGGGA GCGCGGATGT AACCTTTCAA AGCGCTAATA ACTTAGTGTT GAATAAAGCT 1140
AACATAGAAG CTCAAGCCAC AGACAATATC TTTAATCTTT TGGGTCAAGA AGGGATTGAT 1200
AAAATCTTTA ATCAGGGGAA TTTAGCGAAT GTTCTTAGTC AAATGGCTAT GGAAAAAATC 1260
AAGCAAGCCG GCGGTTTAGG GAACTTTATA GAAAACGCTC TAAGCCCTTT GAGTAAGGAA 1320
TTACCCGCTA GCTTGCAAGA TGAAACCTTA GGCCAACTTA TAGGTCAAAA TAACTTAGAT 1380
GATTTATTGA ATAATAGTGG AGTCATGAAT GAAATCCAAA ACATTATCAG TCAAAAACTA 1440
AGCATTTTTG GCAATTTTGT TACCCCATCC ATCATAGAAA ACTACCTTGC TAAGCAGTCT 1500
TTAAAAAGCA TGCTAGACGA TAAAGGGCTT TTGAATTTTA TCGGTGGGTA TATAGACGCT 1560
TCTGAATTAA GCTCTATTTT AGGCGTGATT TTAAAGGATA TTACTAACCC CCCTACAAGC 1620
CTGCAAAAAG ACATTGGTGT GGTAGCGAAC GACTTGTTGA ACGAGTTTTT AGGACAAGAT 1680
GTTGTCAAAA AGCTAGAAAG TCAAGGCTTG GTGAGTAATA TCATCAATAA TGTTATTTCT 1740
CAAGGCGGGT TGAGCGGCGT TTATAATCAA GGTTTAGGGA GCGTGTTGCC GCCCTCTTTA 1800
CAAAACGCGC TCAAAGAAAA CGATTTAGGC ACTCTTTTAT CGCCTAGAGG CTTGCATGAT 1860
TTTTGGCAAA AAGGGTATTT TAACTTTTTA AGCAATGGCT ATGTTTTTGT CAATAACAGC 1920
TCTTTTAGTA ACGCTACTGG GGGTAGTTTG AATTTTGTCG CCAACAAGTC TATTATCTTT 1980
AATGGCGATA ATACGATTGA CTTTAGCAAG TATCAAGGCG CATTGATTTT TGCTTCTAAT 2040
GGTGTTTCTA ATATCAATAT CACCACCCTA AACGCCACTA ATGGCTTAAG CCTTAATGCG 2100
GGTTTGAATA ATGTGAGCGT TCAAAAAGGA GAAATTTGTA TCAATTTAGC CAATTGCCCT 2160
ACAACCAAAA ACAGCTCTCC TGCAAACTCT AGCGTAACCC CCACTAATGA GTCTTTAAGC 2220
GTGCACGCTA ATAATTTCAC TTTCTTAGGC ACAATCATCT CTAATGGGGC TATTGATTTG 2280
TCTCAAGTAA CAAATAATAG CGTTATAGGC ACGCTCAATC TCAATGAAAA TGCGACCTTG 2340
CAAGCTAATA ATTTAACGAT CACCAACGCT TTTAACAACG CCTCTAACTC TACGGCTAAT 2400
ATTGATGGTA ATTTCACCTT AAACCAACAA GCGACTTTAA GCACTAACGC TAGTGGTTTG 2460
AATGTCATGG GGAATTTTAA TAGCTATGGC GATTTGGTGT TTAACCTCAG TCATTCAGTT 2520
AGTCATGCTA TTATCAATAC TCAAGGCACA GCGACGATCA TGGCCAATAA TAACCCTTTG 2580
ATCCAATTCA ACGCTTCTTC AAAAGAAGTG GGTACTTACA CGCTGATTGA TAGCGCTAAA 2640
GCCATTTATT ACGGGTATAA CAACCAAATC ACAGGAGGCA GTAGCCTGGA TAATTACCTT 2700
AAGCTTTATG CGCTCATTGA TATTAATGGC AAGCACATGG TGATGACTGA CAACGGCTTA 2760
ACCTATAACG GGCAAGCCGT GAGCGTTAAA GATGGCGGTT TAGTTGTAGG CTTTAAGGAC 2820
TCTCAAAATC AATACATTTA CACTTCCATT CTTTATAATA AAGTGAAAAT CGCTGTTTCT 2880
AATGATCCTA TCAATAACCC ACAAGCCCCC ACTTTAAAAC AATATATCGC TCAAATTCAG 2940
GGCGTTCAAA GCGTGGATAG CATCGATCAA GCTGGGGGAA ATCAAGCGAT TAATTGGCTC 3000
AATAAAATCT TTGAAACTAA AGGAAGCCCT TTATTCGCTC CCTATTATCT AGAGAGCCAC 3060
TCCACAAAAG ATTTAACCAC GATCGCTGGA GATATTGCTA ACACTTTAGA AGTCATCGCT 3120
AACCCTAATT TTAAAAATGA CGCCACTAAT ATTTTACAGA TCAACACCTA CACGCAGCAA 3180
ATGAGTCGTT TAGCCAAGCT CTCTGACACT TCAACTTTCG CCCGTTCTGA TTTCTTAGAA 3240
CGCTTAGAAG CCCTTAAAAA CAAGCGATTC GCTGATGCGA TCCCTAACGC TATGGATGTG 3300
ATTTTAAAAT ACTCTCAAAG GAATAGAGTT AAAAATAATG TGTGGGCGAC AGGAGTTGGA 3360
GGGGCTAGTT TCATTAGTGG AGGTACTGGA ACTTTATATG GTATCAATGT AGGGTATGAT 3420
AGGTTTATTA AGGGCGTGAT TGTGGGAGGT TATGCCGCTT ATGGGTATAG CGGGTTCCAT 3480
GCAAACATCA CTCAATCAGG CTCTAGCAAT GTCAATGTGG GCGTTTATAG CCGAGCGTTT 3540
ATCAAAAGAA GCGAGCTAAC CATGAGCTTG AATGAGACTT GGGGATACAA TAAAACTTTC 3600
ATCAACTCCT ATGACCCCCT ACTCTCAATC ATCAATCAGT CTTACAGATA CGACACTTGG 3660
ACGACTGACG CTAAAATCAA TTATGGCTAT GATTTCATGT TTAAAGATAA AAGCGTTATT 3720
TTTAAACCCC AAGTAGGCTT AAGCTATTAT TACATTGGTT TGTCTGGTTT AAGGGGCATT 3780
ATGGATGATC CTATTTACAA CCAATTCAGA GCCAATGCTG ACCCTAATAA AAAATCCGTT 3840
CTAACGATCA ATTTTGCCCT AGAAAGTCGG CATTATTTCA ATAAAAACTC TTATTATTTT 3900
GTGATTGCGG ATGTGGGCAG AGACTTATTC ATTAATTCTA TGGGGGATAA AATGGTGCGT 3960

```
TTCATCGGTA ATAACACCCT AAGCTATAGA GATGGTGGCA GATACAACAC TTTTGCTAGC  4020
ATTATCACAG GCGGGGAGAT AAGATTGTTC AAAACCTTTT ATGTGAATGC GGGCATAGGG  4080
GCTAGGTTTG GGCTTGATTA TAAAGATATT AATATTACCG GAAATATTGG TATGCGCTAT  4140
GCTTTTTAA                                                          4149
```

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 789 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...789

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```
ATGAAAAAAA TTGGTTTGAG CTTGTGTTTG GTTTTGAGTT TGGGTTTTTT AAAAGCCCAT    60
GAAGTGAGCG CTGAAGAGAT TGCGGATATT TTCTACAAAC TCAACGCCAA AGAGCCTAAA   120
ATGAAAATCA ACCACACGAA GGGGTTTTGC GCTAAAGGCG TGTTCCTCCC TAACCCGCAA   180
GCAAGAGAGG ATTTAGAGGT GCCACTACTC AATGAAAAAG AAATCCCTGC GTCTGTAAGG   240
TATTCTTTAG GGGGCGTGGC GATGGACGAT AAAAGCAAGG TTAGGGGAAT GGCGTTAAAA   300
CTAGAAAATC AAAACGCTAG TTGGACAATG GTGATGCTCA ATACAGAAAT CAATTTTGCC   360
AAAAACCCTG AAGAATTCGC CCAATTTTTT GAAATGAGAC TTCCTAAAAA TGGCAAGGTA   420
GATGAAGCAA GAATCAAAAA GCTTTACGAA GAAGTCCCCT CTTATAGGAA TTTTGCCGCC   480
TATATGAAAA CGATAGGGAT TAGCTCAAGC GTGGCTAATA CGCCTTATTA TAGCGTGCAT   540
GCGTTCAAGT TTAAAGATAA GAAAGAAAAA TTATTGCCTG CGAGGTGGAA ATTTGTGCCT   600
AAAGAGGGCG TTAAATACTT AAATCCTCAA GAATTAAAGC AAAAAGATTC AAATTATCTG   660
CTCTCTTCAT TCCAACAACA CCTTAAAAAT AAACCCATAG AATACCAAAT GTATTTGGTG   720
TTTGCGAATC AAAATGATGC CACCAACGAC ACGACCGCGC TTTGGAAAGG CAGCATAAGG   780
AATTATTAG                                                          789
```

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 741 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...741

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

ATGAAACAAT TTAAAAAGAA ACCAAAAAAG ATAAAACGAT CGCATCAAAA TCAAAAAACA 60
ATCTTAAAGC GTCCTTTATG GCTTATGCCT TTACTGATTG GCGGGTTTGC TAGTGGGGTG 120
TATGCGGATG GAACAGACAT TTTGGGGCTT AGTTGGGGGG AAAAAAGCCA AAAGGTATGC 180
GTGCATCGTC CATGGTATGC TATATGGAGT TGCGATAAAT GGGAGGAAAA AACACAACAA 240
TTTACAGGAA ACCAACTCAT CACAAAAACT TGGGCAGGGG GTAATGCGGC TAACTACTAC 300
CACTCTCAAA ACAACCAAGA CATCACAGCC AATTTAAAAA ATGATAACGG CACTTATTTT 360
TTAAGCGGTC TGTATAACTA CACCGGAGGG GAATATAATG GGGGGAATTT AGACATTGAA 420
TTAGGCAGTA ACGCTACTTT TAATCTAGGT GCGAGTAGTG GGAATAGCTT CACTTCTTGG 480
TATCCTAATG GGCATACTGA TGTTACTTTT AGCGCTGGGA CTATCAATGT GAATAACAGC 540
GTAGAAGTGG GCAATCGTGT GGGATCGGGA GCTGGCACGC ACACCGGCAC AGCCACTTTA 600
AACTTGAACG CTAATAAGGT TACTATCAAT TCCAATATCA GCGCGTATAA AACTTCGCAA 660
GTGAATGTAG GCAATGCTAA CAGCGTTATT ACCATTAATT CGGTTTCTTT AAATGGGGAA 720
TACTTGCAGT TCTTTAGCTA G 741

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 738 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...738

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

ATGATAAAAA AGACCCTTGC ATCGGTTTTA TTAGGATTGA GTTTGATGAG TGTGTTAAAT 60
GCCAAAGAAT GCGTTTCGCC CATAACAAGA AGCGTTAAGT ATCATCAGCA AAGTGCTGAG 120
ATCAGAGCCT TGCAATTACA AAGTTACAAA ATGGCGAAAA TGGCGCTAGA CAATAACCTT 180
AAGCTCGTTA AAGACAAAAA GCCAGCCGTC ATCTTGGATT TAGATGAAAC CGTTTTGAAC 240
ACTTTTGATT ATGCGGGCTA TTTAGTCAAA AACTGCATTA AATACACCCC AGAAACTTGG 300
GATAAATTTG AAAAAGAAGG CTCTCTTACG CTCATTCCTG GAGCGCTAGA CTTTTTAGAA 360
TACGCTAATT CTAAGGGCGT TAAGATTTTT TACATTTCTA ACCGCACCCA AAAAAATAAG 420

GCATTCACTT TAAAAACGCT CAAAAGCTTT AAGCTCCCCC AAGTGAGTGA AGAATCCGTT 480
TTGTTAAAGG AAAAAGGCAA GCCTAAAGCC GTTAGGCGGG AGTTAGTCGC TAAGGATTAT 540
GCGATTGTTT TACAAGTGGG CGACACTTTG CATGATTTTG ACGCCATTTT TGCTAAAGAC 600
GCTAAAAACA GCCAAGAACA ACAAGCCAAA GTCTTGCAAA ACGCTCAAAA ATTCGGCACA 660
GAATGGATCA TTTTACCCAA CTCTCTTTAT GGCACATGGG AAGATGGGCC TATAAAAGCA 720
TGGCAAAATA AAAAATAA 738

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 867 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...867

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

TTGTGGTGTT TAAAAACCCC TATCATAGGG CATGGCATGA AGAAAAAAGC AAAAGTCTTT 60
TGGTGTTGTT TTAAAATGAT TCGTTGGTTG TATTTGGCGG TCTTTTTTTT GTTGAGCGTA 120
TCAGACGCTA AAGAAATCGC TATGCAACGA TTTGACAAAC AAAACCATAA GATTTTTGAA 180
ATCCTTGCGG ATAAAGTGAG CGCCAAAGAC AATGTGATAA CCGCCTCAGG GAATGCGATC 240
CTATTGAATT ATGACGTGTA TATTCTAGCG GATAAGGTGC GTTATGACAC CAAGACTAAA 300
GAAGCGTTAT TAGAAGGCAA TATTAAGGTT TATAGGGGCG AGGGCTTGCT CGTTAAAACC 360
GATTATGTGA AATTGAGTTT GAACGAAAAA TATGAGATCA TTTTCCCCTT TTATGTCCAA 420
GACAGCGTGA GCGGGATTTG GGTGAGCGCG GATATTGCTA GCGGGAAGGA TCAAAAATAT 480
AAGATTAAAA ACATGAGCGC TTCAGGGTGC AGCATTGACA ACCCCATTTG GCATGTCAAT 540
GCGACTTCAG GCTCATTTAA CATGCAAAAA TCGCATTTGT CAATGTGGAA TCCTAAGATT 600
TATGTCGGCG ATATTCCTGT ATTGTATTTG CCCTATATTT TCATGTCCAC GAGCAATAAA 660
AGAACTACCG GGTTTTTATA CCCTGAGTTT GGCACTTCCA ACTTAGACGG CTTTATTTAT 720
TTGCAACCCT TTTATTTAGC CCCCAAAAAC TCATGGGATA TGACCTTTAC CCCACAAATC 780
CGTTACAAAA GGGGTTTTGG CTTGAATTTT GAAGCGCGCT ACATCAACTC TAAGACGCAG 840
GTTTTTATTC AATGCGCGCT ATTTTAG 867

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 387 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...387

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

TTGATGTTTA AAAAAATGTG TTTGAGCCTG CTAATGATAA GCGGTGTTTG TGTGGGGGCA 60
AAGGATTTGG ATTTCAAGCT GGATTATCGC GCGACTGGGG GGAAATTCAT GGGGAAAATG 120
ACGGACTCTA GTCTTTTAAG TATCACTTCT ATGAACGATG AACCGGTGGT GATTAAAAAC 180
CTTATTGTCA ATAGGGGAAA TTCATGCGAA GCGACTAAAA AAGTAGAACC CAAATTTGGC 240
GATAAGTTTA AAAAAGAAAA ACTCTTTGAT CATGAATTAA AATACTCGCA ACAGATATTT 300
TACCGCCTGG ATTGCAAGCC TAACCAATTG TTAGAAGTTA AAATCATCAC GGACAAGGGC 360
GAATATTACC ATAAATTTTC CAAATAG 387

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 510 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...510

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

ATGCAAGCGT TAAAATCATT GCTTGAAGTG ATTACAAAAC TCCAGAATCT AGGCGGCTAT 60
TTGATGCATA TAGCTATTTT CATCATTTTT ATTTGGATTG GAGGGCTTAA GTTTGTGCCT 120
TACGAAGCTG AAGGGATCGC CCCTTTTGTG GCCAACTCCC CTTTCTTTTC TTTCATGTAT 180
AAATTTGAAA AACCTGCATA CAAACAACAC AAAATGTCTG AATCCCAATC CATGCAAGAA 240
GAAATGCAAG ATAACCCTAA AATCGTTGAA AACAAAGAAT GGCATAAAGA AAACCGCACT 300
TATTTAGTGG CTGAAGGTTT AGGGATTACG ATCATGATCC TAGGCATTTT GGTGCTTTTG 360
GGGCTTTGGA TGCCTTTAAT GGGCGTAGTT GGGGGCTTGC TTGTCGCTGG AATGACGATC 420
ACCACCCTAT TCTTTTTTAT TCACAACGCC AGAAGTGTTT GTCAATCAGC ATTTCCCATG 480
GCTTTCTGGG GCTGGAAGGC TAGTGGTTAA 510

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1464 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1464

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

```
ATGATTGAAT GGATGCAAAA TCATAGAAAG TATTTAGTGG TTACGATATG GATAAGCACG     60
ATCGCTTTTA TTGCCGCCGG AATGATAGGT TGGGGGCAAT ACAGCTTTTC TTTAGATAGC    120
GATAGCGCTG CCAAAGTGGG ACAGATTAAG ATTTCTCAAG AAGAATTAGC CCAAGAATAC    180
CGCCGCCTTA AAGACGCCTA TGCTGAGTCT ATCCCTGATT TTAAAGAACT CACCGAAGAT    240
CAAATCAAAG CCATGCATTT AGAAAAAAGC GCGCTAGATT CGCTCATCAA TCAAGCTTTA    300
TTGAGGAATT TCGCTTTAGA TTTAGGGCTT GGTGCTACCA AGCAAGAAGT GGCCAAAGAG    360
ATCAGAAAAA CGAACGTTTT TCAAAAAGAT GGCGTTTTTG ATGAAGAATT GTATAAAAAT    420
ATCTTAAAAC AAAGCCATTA CCGCCCCAAG CATTTTGAAG AAAGCGTTGA AAGGCTTTTA    480
ATCCTTCAAA AAATCAGCGC TCTATTCCCC AAAACCACCA CCCCTTTGGA GCAATCCAGT    540
CTATCGCTTT GGGCAAAATT GCAAGACAAA TTAGACATTC TTATCCTAAA TCCTAATGAT    600
GTTAAAATCT CTCTCAATGA AGAAGAGATG AAAAAATATT ATGAAAACCA TAGAAAGGAT    660
TTTAAAAAGC CCACAAGCTT TAAAACACGC TCTTTATATT TTGACGCTAG TTTAGAAAAA    720
ACTGATTTGA AAGAGTTGGA GGAATACTAC CATAAAAACA AGGTGTCTTA TTTGGACAAA    780
GAGGGGAAAT TACAGGATTT TAAAAGCGTT CAAGAGCAAG TCAAGCATGA TTTAAACATG    840
CAAAAGGCGA ATGAAAAAGC CTTAAGGAGC TATATCGCTC TAAAAAAGGG GAACGCACAA    900
AACTACACCA CGCAAGATTT TGAAAAAAAC AACTCCCCCT ATACTGCTGA AATCACGCAA    960
AAACTCACCG CTCTCAAGCC CCTTGAAGTC CTAAAACCAG AGCCTTTTAA AGATGGTTTT   1020
ATCGTGGTGC AGCTTGTCTC TCAAATTAAA GACGAATTGC AAAATTTTGA TGAAGCCAAA   1080
AGCGCTCTTA AAACCCGTCT GACTCAAGAA AAAACCCTTA TGGCGTTGCA AACTTTAGCT   1140
AAAGAAAAGC TTAAGGATTT TAAAGGGAAA AGCGTGGGTT ATGTAAGCCC TAATTTTGGA   1200
GGCACTATCA GTGAACTTAA CCAAGAAGAG AGCGCGAAGT TTATCAACAC CCTTTTTAAC   1260
CGCCAGGAAA AAAAAGGGTT TGTAACCATA GGTAATAAAG TGGTGCTTTA TCAAATCACA   1320
GAGCAAAATT TCAATCACCC CTTTAGTGCA GAAGAAAACC AATACATGCA GCGTTTAGTC   1380
AATAACACTA AAACGGATTT TTTTGATAAA GCGTTGATAG AAGAATTGAA AAAACGCTAT   1440
AAGATAGTCA AATACATTCA ATAA                                          1464
```

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 429 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...429

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATGAAAACGA ACTTTTATAA AATTAAATTA CTATTTGCTT GGTGTCTTAT CATTGGCATG 60
TTTAACGCTC CGCTTAACGC TGACCAAAAC ACGGATATAA AAGATATTAG TCCTGAAGAT 120
ATGGCGCTAA ATAGCGTGGG GCTTGTTTCT AGAGATCAGC TAAAAATAGA GATCCCTAAA 180
GAAACCCTAG AGCAAAAAGT GGCCATACTC AATGACTATA ATGATAAGAA TGTTAATATC 240
AAGTTTGACG ACATAAGTTT AGGGAGTTTC CAACCTAATG ATAATCTAGG TATCAATGCG 300
ATGTGGGGCA TTCAAAATCT TCTCATGAGC CAAATGATGA GCAATTACGG TCCAAACAAT 360
TCTTTCATGT ATGGCTATGC GCCAACATAC TCAGATTCAT CGTTTTTACC ACCGATCTTA 420
GGGTATTAA 429

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 627 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...627

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

TTGATCAACA ATAATAATAA CAATAAAAAA CTGAGAGGCT TTTTTTTGAA AGTTCTCTTA 60
AGTCTCGTTG TTTTCAGTTC GTATGGGTCA GCAAATGACG ATAAAGAAGC CAAAAAAGAA 120
GCGCTAGAAA AAGAAAAAAA CACTCCCAAT GGGCTTGTTT ATACGAATTT AGATTTTGAT 180
AGTTTTAAAG CGACTATCAA AAATTTGAAA GACAAGAAAG TAACTTTCAA AGAAGTCAAT 240
CCCGATATTA TCAAAGATGA AGTTTTTGAC TTCGTGATTG TCAATAGAGT CCTTAAAAAA 300
ATAAAGGATT TGAAGCATTA CGATCCAGTT ATTGAAAAAA TCTTTGATGA AAAGGGTAAA 360

GAAATGGGAT TGAATGTAGA ATTACAGATC AATCCTGAAG TGAAAGACTT TTTTACTTTC 420
AAAAGCATCA GCACGACCAA CAAACAACGC TGCTTTCTAT CATTGCACGG AGAAACAAGA 480
GAAATTTTAT GCGATGATAA GCTATATAAT GTTTTATTGG CCGTATTCAA TTCTTATGAT 540
CCTAATGATC TTTTGAAACA CATTAGCACC ATAGAGTCTC TCAAAAAAAT CTTTTATACG 600
ATTACATGTG AAGCGGTATA TCTATAA 627

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 738 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...738

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

ATGGCAGGCA CACAAGCTAT ATATGAATCA TCTTCTGCAG GATTCTTATC GCAAGTCTCC 60
TCAATCATCT CAAGCACAAG TGGTGTCGCA GGGCCATTTG CAGGAATAGT AGCGGGCGCT 120
ATGACAGCAG CGATTATTCC TATTGTTGTG GGATTTACTA ATCCGCAAAT GACCGCTATC 180
ATGACCCAAT ACAATCAAAG CATCGCTGAA GCTGTAAGCG TGCCTATGAA AGCCGCTAAC 240
CAACAATACA ACCAATTGTA TCAAGGTTTT AACGATCAAA GCATGGCTGT GGGGAACAAT 300
ATCTTAAATA TCAGCAAATT AACAGGGGAA TTTAACGCGC AAGGCAACAC GCAAAGCGCG 360
CAAATTAGTG CTGTCAATAG TCAGATTGCA AGCATTTTAG CGAGTAACAC TACCCCTAAA 420
AATCCTAGCG CTATTGAAGC TTATGCGACG AATCAAATCG CTGTTCCTAG CGTGCCAACA 480
ACGGTTGAAA TGATGAGCGG TATATTAGGC AATATTACAA GCGCAGCACC AAAATACGCC 540
CTAGCTCTAC AAGAGCAACT GCGTTCTCAA GCAAGCAACA GCTCAATGAA TGATACAGCC 600
GATTCCCTTG ATAGCTGTAC CGCTTTAGGC GCACTTGTTG GCTCATCAAA AGTGTTTTTC 660
AGTTGCATGC AAATTTCTAT GACTCCTATG AGTGTTTCTA TGCCCACTGT TATGCCAAAT 720
ACCAGCGGTT GCCACTAA 738

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1104 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1104

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

ATGATTAAAA GCGTAGAGAT TGAAAATTAC AAAAATTTTG AGCACCTTAA AATGGAAAAT 60
TTTAAACTCA TCAACTTTTT TACCGGTCAA AACGATGCGG GTAAAACCAA TCTTTTAGAA 120
GCTCTTTATA CCAACACAGG CCTTTGTGAT CCTACTGCCA ATCAAGTCAG TCTTCCTCCT 180
GAACATGCCG TGAATATTAG TGAATTCAGA AAAATCAAAC TCGATGCCGA CAACCTAAAA 240
ACCTTTTTTT ATCAAGGAAA CACCGCTAAT CCCATTAGTA TCCGCACTGA ATTTGAACAT 300
GCTACTATCC CTCTTACTAT CCAATACCCC ACACAAACCA GTTACAGCAA AGACATCAAT 360
TTGAATAGCG ATGATGCTCA TATGACAAAC CTTATAAACA CAACAATAAC GAAGCCACAG 420
CTCCAATTTT CCTACAATCC ATCCCTTTCC CCCATGACAA TGACTTATGA ATTTGAAAGG 480
CAAAACCTAG GTTTAATCCA TTCTAATTTA GATAAAATCG CTCAAACCTA TAAAGAAAAT 540
GCGATGTTTA TTCCTATAGA ATTATCTATT GTTAATTCTC TTAAAGCATT GGAAAATTTA 600
CAATTAGCAA GCAAAGAAAA AGAATTGATT GAAATCCTAC AATGTTTCAA CCCTAATATT 660
TTAAATGCTA ATACAATAAG AAAGTCTGTC TATATCCAAA TCAAAGATGA AAACACACCG 720
CTAGAAGAAA GTCCCAAAAG GCTTTTAAAT TTGTTTGGTT GGGGTTTTAT CAAATTCTTT 780
ATTATGGTGA GCATTCTTAT AGACAATCGT GTCAAGTATC TTTTTATTGA TGAAATAGAA 840
AGCGGTTTGC ACCATACAAA AATGCAAGAG TTTTTAAAAG CTCTGTTTAA GTTAGCTCAA 900
AAATTACAGA TTCAAATTTT TGCCACCACG CACAATAAGG AATTTTTATT AAACGCCATC 960
AACACGATAT CCGATAATGA AACGGGAGTT TTTAAAGACA TAGCCTTGTT TGAGCTTGAA 1020
AAAGAAAGCG CTTCTGGCTT TATCAGACAC AGCTATTCTA TGCTAGAAAA AGCGCTTTAT 1080
AGGGGTATGG AGGTTAGAGG CTGA 1104

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1230 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1230

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

ATGTCCTTGA TTAGAGTGAA TGGGGAAGCT TTTAAACTCT CTTTGGAAAG TTTAGAAGAA 60

GATCCTTTTG AAACTAAAGA AACGCTAGAA ACGCTAGAAA CGCTTATCAA ACAAACGAGC 120
GTTGTTTTAT TGGCCGCTGG GGAGTCTAAG CGTTTTTCTC GTGCGATTAA AAAGCAGTGG 180
CTACGCTCTC ACCACACCCC CTTATGGCTC AGCGTGTATG AAAGCTTTAA AGAAGCCCTA 240
GACTTTAAGG AAGTCATTCT AGTTGTAAGC GAATTGGATT ATGTTTATAT CCAACGCCAT 300
TACCCCAAAA TCAAGCTTGT AAAAGGCGGG GCATCAAGGC AAGAATCCGT GCGTAACGCT 360
TTGAAAGTAA TTGATAGCAC TTACACGATC ACCAGCGATG TGGCTAGGGG TTTAGCGAAT 420
ATGGAAGCGC TTAAAAGCTT GTTTTTAACC CTCCAACAAA CGAGCCATTA TTGCATCGCC 480
CCTTACTTGC CTTGCTATGA CACAGCGATC TATTATAACG AGGCTTTAGA TAGAGAAGCG 540
ATCAAACTCA TTCAAACCCC GCAATTAAGC CACACCAAAA CGCTCCAATC AGCCCTAAAC 600
CAAGGGGGTT TTAAAGATGA AAGCAGCGCG ATTTTACAAG CTTTCCCTAA CTCTGTGAGC 660
TATATTGAAG GCAGTAAGGA TTTGCACAAA CTCACCACAA GCGGCGATTT AAAGTTTTTT 720
ACGCCTTTTT TTAACCCAGC AAAGGACACT TTTATAGGCA TGGGTTTTGA TACGCATGCG 780
TTCATTAAAG ATAAGCCTAT GGTTTTAGGG GGGGTTGTTT TGGATTGCGA GTTTGGGTTA 840
AAGGCTCATA GCGATGGCGA TGCTTTATTG CATGCGGTTA TTGATGCGAT TTTAGGAGCG 900
ATTAAAGGGG GGGATATTGG CGAATGGTTC CCTGATAATG ACCCCAAATA CAAAAACGCC 960
TCTTCTAAAG AGCTTTTAAA AATCGTGTTG GATTTTTCTC AAAGCATTGG GTTTGAATTG 1020
CTTGAAATGG GAGCGACCAT CTTTAGCGAA ATCCCTAAAA TCACTCCTTA CAAACCGGCG 1080
ATTTTAGAGA ATTTGAGCCA ACTTTTGGGT TTAGAAAAAT CTCAAATCAG CTTGAAAGCC 1140
ACTACAATGG AAAAAATGGG GTTCATTGGC AAACAAGAAG GGCTGTTAGT CCAAGCGCAT 1200
GTGAGCATGC GTTATAAACA AAAACTTTAA 1230

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 813 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...813

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

ATGAAAAAGT TTGTAGCTTT AGGGCTTCTA TCCGCGGTTT TAAGCTCTTC GTTGTTAGCC 60
GAAGGTGATG GTGTTTATAT AGGGACTAAT TATCAGCTTG GACAAGCCCG TTTGAATAGC 120
AATATTTATA ATACAGGGGA TTGCACAGGG AGTGTTGTAG GTTGCCCCCC AGGTCTTACC 180
GCTAATAAGC ATAATCCAGG AGGCACCAAT ATCAATTGGC ACTCCAAATA CGCTAATGGG 240
GCTTTGAATG GTTTTGGGTT GAATGTGGGT TATAAGAAAT TCTTCCAATT CAAGTCGCTA 300
GATATGACAA GCAAGTGGTT TGGTTTTAGA GTGTATGGGC TTTTTGATTA CGGGCATGCC 360
GATTTAGGTA AACAAGTTTA TGCACCTAAT AAAATCCAGT TGGATATGGT CTCTTGGGGT 420
GTGGGGAGCG ATTTGTTAGC TGATATTATT GATAAAGACA ACGCTTCTTT TGGTATTTTT 480
GGTGGGGTCG CTATCGGCGG TAACACTTGG AAAAGCTCTG CAGCAAACTA TTGGAAAGAG 540
CAAATCATTG AAGCCAAAGG TCCTGATGTT TGTACCCCTA CTTATTGTAA CCCTAATGCC 600
CCTTATAGCA CCAACACTTC AACCGTCGCT TTTCAAGTGT GGTTGAATTT TGGGGTGAGA 660

GCCAATATCT ACAAGCATAA TGGCGTGGAA TTTGGCGTGA GAGTGCCGCT ACTCATCAAT 720
AAATTTTTGA GCGCGGGTCC TAACGCTACT AACCTTTATT ACCATTTGAA ACGGGATTAT 780
TCGCTTTATT TGGGGTATAA CTACACTTTT TAA 813

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1317 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1317

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

ATGGCTTACA AACCTAACAA AAAGAAGTTA AAAGAATTAA GAGAGCAACC GAATTTATTT 60
AGCATCTTAG ATAAGGGCGA TGTTGCAACA AACAATCCTG TTGAAGAGTC AGACAAGGCC 120
AATAAAATAC AAGAGCCACT CCCTTATGTC GTGAAAACGC AAATCAATAA AGCAAGCATG 180
ATTTCTAGAG ATCCTATTGA ATGGGCAAAG TATTTAAGCT TTGAAAAACG AGTCTATAAG 240
GATAATAGTA AAGAAGATGT CAATTTCTTT GCCAATGGTG AGATAAAAGA AAGTTCTCGT 300
GTTTATGAAG CGAATAAAGA AGGGTTTGAA AGGCGCATCA CTAAAAGATA CGATCTGATT 360
GATAGAAATA TTGATAGAAA TAGAGAATTT TTTATAAAAG AAATTGAAAT TCTAACCCAC 420
ACAAACAGCT TAAAAGAATT GAAAGAGCAA GGGTTAGAAA TCCAATTGAC CCACCATAAT 480
GAAACGCATA AGAAAGCCTT AGAAAATGGC AATGAAATCG TTAAAGAATA CGACCATCTT 540
AAAGATATTT ACCAAGAAGT AGAAAGAACA AAAGATGGTG GATTGGTAAG AGAAATAATC 600
CCCAGTATTT CTAGCGCTGA GTATTTCAAG CTTTACAACA AACTGCCTTT TGAATCAATA 660
AACAATGAAA ATACCAAACT GAATACTAAC GACAATGAAG AAGTTAAAAA ACTAGAATTT 720
GAATTAGCTA AAGAAGTGCA TATTTTAATC CTAGAGCAAC AATTGCTTTC AGCAACAAAT 780
TATTATTCTT GGATAGATAA AGATGATAAT GCGAATTTTG CTTGGAAAAT GCATAGGCTT 840
ATCAATGAAA ATAAACTCAA AGAAAACCAT CTCAGCGCCA ATAACGCTAA TAAGATTAAG 900
CAATTTTTCT TTAATAATGG TTCTATTTTA GGCTGGACTA AAGAAGAACA AAGCGCTATA 960
CAAGAAAACA GAGATTATTC TTTAAGAAGC GCTCTTTTAA GTTTAGAAGA AATCGCTCAA 1020
GCAAAAATTG AATTGCAAAA ATACTATGAA AGCGTTTATG TTAATGGTGA TGGGAATAAA 1080
AGAGAAATCA AGCCTTTTAA AGAAATTTTA AGAGACACCA ACAATTTTGA AAAAGCTTAT 1140
AAGGAGCGTT ATGACAAATT GGTAAGCTTG AGTGCAGCAA TCATTCAAGC TAAAGAGGGT 1200
GGTAATGAGC GACCAAATTC TAGTGCAAAT AACAATAACC CTATTAAAAA TACAATAGAG 1260
ACTAATACTT CTAACAATAT TATTCAAAAT AATGATAATA TAATCATCCA AATTTAA 1317

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 648 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...648

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

```
ATGCAAGCGT TAAAATCATT GCTTGAAGTG ATTACAAAAC TCCAGAATCT AGGCGGCTAT    60
TTGATGCATA TAGCTATTTT CATCATTTTT ATTTGGATTG GAGGGCTTAA GTTTGTGCCT   120
TACGAAGCTG AAGGGATCGC CCCTTTTGTG GCCAACTCCC CTTTCTTTTC TTTCATGTAT   180
AAATTTGAAA AACCTGCATA CAAACAACAC AAAATGTCTG AATCCCAATC CATGCAAGAA   240
GAAATGCAAG ATAACCCTAA AATCGTTGAA AACAAAGAAT GGCATAAAGA AAACCGCACT   300
TATTTAGTGG CTGAAGGTTT AGGGATTACG ATCATGATCC TAGGCATTTT GGTGCTTTTG   360
GGGCTTTGGA TGCCTTTAAT GGGCGTAGTT GGGGGCTTGC TTGTCGCTGG AATGACGATC   420
ACCACCCTAT CTTTTTTATT CACAACGCCA GAAGTGTTTG TCAATCAGCA TTTCCCATGG   480
CTTTCTGGGG CTGGAAGGCT AGTGGTTAAA GACTTGGCGT TATTTGCTGG AGGCTTGTTT   540
GTGGCCGGAT TTGATGCGAA ACGCTATTTG GAGGGTAAAG GGTTTTGCTT GATGGACCGC   600
TCATCGGTAG GGATTAAAAC TAAATGCTCT AGCGGGTGTT GCTCTTAA               648
```

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 186 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...186

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

```
Met Ile Lys Arg Ile Ala Cys Ile Leu Ser Leu Ser Ala Ser Leu Ala
1               5                  10                  15
Leu Ala Gly Glu Val Asn Gly Phe Phe Met Gly Ala Gly Tyr Gln Gln
            20                  25                  30
Gly Arg Tyr Gly Pro Tyr Asn Ser Asn Tyr Ser Asp Trp Arg His Gly
```

```
          35                  40                  45
Asn Asp Leu Tyr Gly Leu Asn Phe Lys Leu Gly Phe Val Gly Phe Ala
      50                  55                  60
Asn Lys Trp Phe Gly Ala Arg Val Tyr Gly Phe Leu Asp Trp Phe Asn
65                    70                  75                  80
Thr Ser Gly Thr Glu His Thr Lys Thr Asn Leu Leu Thr Tyr Gly Gly
                  85                  90                  95
Gly Gly Asp Leu Ile Val Asn Leu Ile Pro Leu Asp Lys Phe Ala Leu
             100                 105                 110
Gly Leu Ile Gly Gly Val Gln Leu Ala Gly Asn Thr Trp Met Phe Pro
         115                 120                 125
Tyr Asp Val Asn Gln Thr Arg Phe Gln Phe Leu Trp Asn Leu Gly Gly
     130                 135                 140
Arg Met Arg Val Gly Asp Arg Ser Ala Phe Glu Ala Gly Val Lys Phe
145                   150                 155                 160
Pro Met Val Asn Gln Gly Ser Lys Asp Val Gly Leu Ile Arg Tyr Tyr
                  165                 170                 175
Ser Trp Tyr Val Asp Tyr Val Phe Thr Phe
             180                 185
```

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 116 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...116

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

```
Leu Met Arg Ile Ile Ile Arg Leu Leu Ser Phe Lys Met Asn Ala Phe
1               5                  10                  15
Leu Lys Leu Ala Leu Ala Ser Leu Met Gly Gly Leu Trp Tyr Ala Phe
            20                  25                  30
Asn Gly Glu Gly Ser Glu Ile Val Ala Ile Gly Ile Phe Val Leu Ile
        35                  40                  45
Leu Phe Val Phe Phe Ile Arg Pro Val Ser Phe Gln Asp Pro Glu Lys
    50                  55                  60
Arg Glu Glu Tyr Ile Glu Arg Leu Lys Lys Asn His Glu Arg Lys Met
65                  70                  75                  80
Ile Leu Gln Asp Lys Gln Lys Glu Glu Gln Met Arg Leu Tyr Gln Ala
                85                  90                  95
Lys Lys Glu Arg Glu Ser Arg Gln Lys Gln Asp Leu Lys Glu Gln Met
            100                 105                 110
Lys Lys Tyr Ser
```

115

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 345 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...345

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

```
Met Val Lys His Tyr Leu Phe Met Ala Val Ser Gln Val Phe Phe Ser
1               5                  10                  15
Phe Phe Leu Val Leu Phe Phe Ile Ser Ser Ile Val Leu Leu Ile Ser
            20                  25                  30
Ile Ala Ser Val Thr Leu Val Ile Lys Val Ser Phe Leu Asp Leu Val
        35                  40                  45
Gln Leu Phe Leu Tyr Ser Leu Pro Gly Thr Ile Phe Phe Ile Leu Pro
    50                  55                  60
Ile Thr Phe Phe Ala Ala Cys Ala Leu Gly Leu Ser Arg Leu Ser Tyr
65                  70                  75                  80
Asp His Glu Leu Leu Val Phe Phe Ser Leu Gly Val Ser Pro Lys Lys
                85                  90                  95
Met Thr Lys Ala Phe Val Pro Leu Ser Leu Leu Val Ser Ala Ile Leu
            100                 105                 110
Leu Ala Phe Ser Leu Ile Leu Ile Pro Thr Ser Lys Ser Ala Tyr Tyr
        115                 120                 125
Gly Phe Leu Arg Gln Lys Lys Asp Lys Ile Asp Ile Asn Ile Arg Ala
    130                 135                 140
Gly Glu Phe Gly Gln Lys Leu Gly Asp Trp Leu Val Tyr Val Asp Lys
145                 150                 155                 160
Thr Glu Asn Asn Ser Tyr Asp Asn Leu Val Leu Phe Ser Asn Lys Ser
                165                 170                 175
Leu Ser Gln Glu Ser Phe Ile Leu Ala Gln Lys Gly Asn Ile Asn Asn
            180                 185                 190
Gln Asn Gly Val Phe Glu Leu Asn Leu Tyr Asn Gly His Ala Tyr Phe
        195                 200                 205
Thr Gln Gly Asp Lys Met Arg Lys Val Asp Phe Glu Glu Leu His Leu
    210                 215                 220
Arg Asn Lys Leu Lys Ser Phe Asn Ser Asn Asp Ala Ala Tyr Leu Gln
225                 230                 235                 240
Gly Thr Asp Tyr Leu Gly Tyr Trp Lys Lys Ala Phe Gly Lys Asn Ala
                245                 250                 255
Asn Lys Asn Gln Lys Arg Arg Phe Ser Gln Ala Ile Leu Val Ser Leu
```

```
                260                 265                 270
Phe Pro Leu Ala Ser Val Phe Leu Ile Pro Leu Phe Gly Ile Ala Asn
        275                 280                 285
Pro Arg Phe Lys Thr Asn Trp Ser Tyr Phe Tyr Val Leu Gly Ala Val
    290                 295                 300
Gly Val Tyr Phe Leu Met Val His Val Ile Ser Thr Asp Leu Phe Leu
305                 310                 315                 320
Met Thr Phe Phe Phe Pro Phe Ile Trp Ala Phe Ile Ser Tyr Leu Leu
                325                 330                 335
Phe Arg Lys Phe Ile Leu Lys Arg Tyr
            340                 345
```

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 276 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...276

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

```
Met Lys Lys Lys Ala Lys Val Phe Trp Cys Cys Phe Lys Met Ile Arg
1               5                   10                  15
Trp Leu Tyr Leu Ala Val Phe Phe Leu Leu Ser Val Ser Asp Ala Lys
            20                  25                  30
Glu Ile Ala Met Gln Arg Phe Asp Lys Gln Asn His Lys Ile Phe Glu
        35                  40                  45
Ile Leu Ala Asp Lys Val Ser Ala Lys Asp Asn Val Ile Thr Ala Ser
    50                  55                  60
Gly Asn Ala Ile Leu Leu Asn Tyr Asp Val Tyr Ile Leu Ala Asp Lys
65                  70                  75                  80
Val Arg Tyr Asp Thr Lys Thr Lys Glu Ala Leu Leu Glu Gly Asn Ile
                85                  90                  95
Lys Val Tyr Arg Gly Glu Gly Leu Leu Val Lys Thr Asp Tyr Val Lys
            100                 105                 110
Leu Ser Leu Asn Glu Lys Tyr Glu Ile Ile Phe Pro Phe Tyr Val Gln
        115                 120                 125
Asp Ser Val Ser Gly Ile Trp Val Ser Ala Asp Ile Ala Ser Gly Lys
    130                 135                 140
Asp Gln Lys Tyr Lys Ile Lys Asn Met Ser Ala Ser Gly Cys Ser Ile
145                 150                 155                 160
Asp Asn Pro Ile Trp His Val Asn Ala Thr Ser Gly Ser Phe Asn Met
                165                 170                 175
Gln Lys Ser His Leu Ser Met Trp Asn Pro Lys Ile Tyr Val Gly Asp
```

```
                180                 185                 190
Ile Pro Val Leu Tyr Leu Pro Tyr Ile Phe Met Ser Thr Ser Asn Lys
        195                 200                 205
Arg Thr Thr Gly Phe Leu Tyr Pro Glu Phe Gly Thr Ser Asn Leu Asp
    210                 215                 220
Gly Phe Ile Tyr Leu Gln Pro Phe Tyr Leu Ala Pro Lys Asn Ser Trp
225                 230                 235                 240
Asp Met Thr Phe Thr Pro Gln Ile Arg Tyr Lys Arg Gly Phe Gly Leu
                245                 250                 255
Asn Phe Glu Ala Arg Tyr Ile Asn Ser Lys Thr Gln Val Phe Ile Gln
            260                 265                 270
Cys Ala Leu Phe
        275
```

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 224 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...224

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

```
Met Ile Arg Leu Lys Gly Leu Asn Lys Thr Leu Lys Thr Ser Leu Leu
1               5                   10                  15
Ala Gly Val Leu Leu Gly Ala Thr Ala Pro Leu Met Ala Lys Pro Leu
            20                  25                  30
Leu Ser Asp Glu Asp Leu Leu Lys Arg Val Lys Leu His Asn Ile Lys
        35                  40                  45
Glu Asp Thr Leu Thr Ser Cys Asn Ala Lys Val Asp Gly Ser Gln Tyr
    50                  55                  60
Leu Asn Ser Gly Trp Asn Leu Ser Lys Glu Phe Pro Gln Glu Tyr Arg
65                  70                  75                  80
Glu Lys Ile Phe Glu Cys Val Glu Glu Glu Lys His Lys Gln Ala Leu
                85                  90                  95
Asn Leu Ile Asn Lys Glu Asp Thr Lys Asp Lys Glu Glu Leu Ala Lys
            100                 105                 110
Lys Ile Lys Glu Ile Lys Glu Lys Ala Lys Val Leu Arg Gln Lys Phe
        115                 120                 125
Met Ala Phe Glu Met Lys Glu His Ser Lys Glu Phe Pro Asn Lys Lys
    130                 135                 140
Gln Leu Gln Thr Met Leu Glu Asn Ala Phe Asp Asn Gly Ala Glu Ser
145                 150                 155                 160
Phe Ile Asp Asp Trp His Glu Arg Phe Gly Gly Ile Ser Arg Glu Asn
```

```
                    165                 170                 175
Thr Tyr Lys Ala Leu Gly Ile Lys Glu Tyr Ser Asp Glu Gly Lys Ile
                180                 185                 190
Leu Pro Leu Ala Lys Glu Val Ile Leu Asp Asn Ile Lys Lys Ile Leu
        195                 200                 205
Lys Lys Ala Leu Met Ile Leu Asp Asn Pro Tyr Leu Leu Trp Leu Val
    210                 215                 220
```

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 429 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...429

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

```
Met Pro Tyr Ala Leu Arg Lys Arg Phe Phe Lys Arg Leu Leu Leu Phe
1               5                   10                  15
Phe Leu Ile Val Cys Met Ile Asn Leu His Ala Lys Ser Tyr Leu Phe
            20                  25                  30
Ser Pro Leu Pro Pro Ala His Gln Gln Ile Ile Lys Thr Glu Pro Cys
        35                  40                  45
Ser Leu Glu Cys Leu Lys Asp Leu Met Leu Gln Asn Gln Ile Phe Ser
    50                  55                  60
Phe Val Ser Gln Tyr Asp Asp Asn Asn Gln Asp Glu Ser Leu Lys Thr
65                  70                  75                  80
Tyr Tyr Lys Asp Ile Leu Asn Lys Leu Asn Pro Val Phe Ile Ala Ser
                85                  90                  95
Gln Thr Pro Ala Lys Glu Ser Tyr Glu Pro Lys Ile Glu Leu Ala Ile
            100                 105                 110
Leu Leu Pro Lys Lys Val Val Gly Arg Tyr Ala Ile Leu Val Met Asn
        115                 120                 125
Thr Leu Leu Ala Tyr Leu Asn Thr Arg Asn Asn Asp Phe Asn Ile Gln
    130                 135                 140
Val Phe Asp Ser Asp Glu Glu Ser Pro Glu Lys Leu Glu Glu Thr Tyr
145                 150                 155                 160
Lys Glu Ile Glu Lys Glu Lys Phe Pro Phe Ile Ile Ala Leu Leu Thr
                165                 170                 175
Lys Glu Gly Val Glu Asn Leu Leu Gln Asn Thr Thr Ile Asn Thr Pro
            180                 185                 190
Thr Tyr Val Pro Thr Val Asn Lys Thr Gln Leu Glu Asn His Thr Glu
        195                 200                 205
Leu Ser Leu Ser Glu Arg Leu Tyr Phe Gly Gly Ile Asp Tyr Lys Glu
```

```
    210                 215                 220
Gln Leu Gly Met Leu Ala Thr Phe Ile Ser Pro Asn Ser Pro Val Ile
225                 230                 235                 240
Glu Tyr Asp Asp Asp Gly Leu Ile Gly Glu Arg Leu Arg Gln Ile Thr
                245                 250                 255
Glu Ser Leu Asn Val Glu Val Lys His Gln Glu Asn Ile Ser Tyr Lys
            260                 265                 270
Gln Ala Thr Ser Phe Ser Lys Asn Phe Arg Lys His Asp Ala Phe Phe
        275                 280                 285
Lys Asn Ser Thr Leu Ile Leu Asn Thr Pro Thr Thr Lys Ser Gly Leu
    290                 295                 300
Ile Leu Ser Gln Ile Gly Leu Leu Glu Tyr Lys Pro Leu Lys Ile Leu
305                 310                 315                 320
Ser Thr Gln Ile Asn Phe Asn Pro Ser Leu Leu Leu Leu Thr Gln Pro
                325                 330                 335
Lys Asp Arg Lys Asn Leu Phe Ile Val Asn Ala Leu Gln Asn Ser Asp
            340                 345                 350
Glu Thr Leu Ile Glu Tyr Ala Ser Leu Leu Glu Ser Asp Leu Arg His
        355                 360                 365
Asp Trp Val Asn Tyr Ser Ser Ala Ile Gly Leu Glu Met Phe Leu Asn
    370                 375                 380
Thr Leu Asp Pro His Phe Lys Lys Ser Phe Gln Glu Ser Leu Glu Asp
385                 390                 395                 400
Asn Gln Val Arg Tyr His Asn Gln Ile Tyr Gln Ala Leu Gly Tyr Ser
                405                 410                 415
Phe Glu Pro Ile Lys Asn Glu Ser Glu Thr Lys Lys Glu
            420                 425
```

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 455 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...455

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

```
Val Leu Lys Phe Gln Lys Leu Pro Leu Leu Phe Val Ser Ile Leu Tyr
1               5                   10                  15
Asn Gln Ser Pro Leu Leu Ala Phe Asp Tyr Lys Phe Ser Gly Val Ala
            20                  25                  30
Glu Ser Val Ser Lys Val Gly Phe Asn His Ser Lys Leu Asn Ser Lys
        35                  40                  45
Glu Gly Ile Phe Pro Thr Ala Thr Phe Val Thr Ala Thr Ile Lys Leu
```

```
    50                  55                  60
Gln Val Asp Ser Asn Leu Leu Pro Lys Asn Ile Glu Lys His Ser Leu
65                  70                  75                  80
Lys Ile Gly Val Gly Gly Ile Leu Gly Ala Leu Ala Tyr Asp Ser Thr
                85                  90                  95
Lys Thr Leu Ile Asp Gln Ala Thr His Gln Ile Tyr Gly Ser Glu Leu
            100                 105                 110
Phe Tyr Leu Ile Gly Arg Trp Trp Gly Phe Leu Gly Asn Ala Pro Trp
        115                 120                 125
Lys Asp Ser Leu Ile Glu Ser Asp Ala His Thr Arg Asn Tyr Val Leu
    130                 135                 140
Tyr Asn Ser Tyr Leu Phe Tyr Ser Tyr Gly Asp Lys Phe His Leu Lys
145                 150                 155                 160
Leu Gly Arg Tyr Leu Ser Asn Met Asp Phe Met Ser Ser Tyr Thr Gln
                165                 170                 175
Gly Phe Glu Leu Asp Tyr Lys Ile Asn Ser Lys Ile Ala Leu Lys Trp
            180                 185                 190
Phe Ser Ser Phe Gly Arg Ala Leu Ala Phe Gly Gln Trp Ile Arg Asp
        195                 200                 205
Trp Tyr Ala Pro Ile Val Thr Glu Asp Gly Arg Lys Glu Val Tyr Asp
    210                 215                 220
Gly Ile His Ala Ala Gln Leu Tyr Phe Ser Ser Lys His Val Gln Val
225                 230                 235                 240
Met Pro Phe Ala Tyr Phe Ser Pro Lys Ile Tyr Gly Ala Pro Gly Val
                245                 250                 255
Lys Ile His Ile Asp Ser Asn Pro Lys Phe Lys Gly Leu Gly Leu Arg
            260                 265                 270
Ala Gln Thr Thr Ile Asn Val Ile Phe Pro Val Tyr Ala Lys Asp Leu
        275                 280                 285
Tyr Asp Val Tyr Trp Arg Asn Ser Lys Ile Gly Glu Trp Gly Ala Ser
    290                 295                 300
Leu Leu Ile His Gln Arg Phe Asp Tyr Asn Glu Phe Asn Phe Gly Phe
305                 310                 315                 320
Gly Tyr Tyr Gln Asn Phe Gly Asn Ala Asn Ala Arg Ile Gly Trp Tyr
                325                 330                 335
Gly Asn Pro Ile Pro Phe Asn Tyr Arg Asn Asn Ser Val Tyr Gly Gly
            340                 345                 350
Val Phe Ser Asn Ala Ile Thr Ala Asp Ala Val Ser Gly Tyr Val Phe
        355                 360                 365
Gly Gly Gly Val Tyr Arg Gly Phe Leu Trp Gly Ile Leu Gly Arg Tyr
    370                 375                 380
Thr Tyr Ala Thr Arg Ala Ser Glu Arg Ser Ile Asn Leu Asn Leu Gly
385                 390                 395                 400
Tyr Lys Trp Gly Ser Phe Ala Arg Val Asp Val Asn Leu Glu Tyr Tyr
                405                 410                 415
Val Val Ser Met His Asn Gly Tyr Arg Leu Asp Tyr Leu Thr Gly Pro
            420                 425                 430
Phe Asn Lys Ala Phe Lys Ala Asp Ala Gln Asp Arg Ser Asn Leu Met
        435                 440                 445
Val Ser Met Lys Phe Phe Phe
    450                 455
```

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 282 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...282

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

```
Met Gly Cys Ser Phe Ile Phe Lys Lys Val Arg Val Tyr Ser Lys Met
1               5                  10                  15
Leu Val Ala Leu Gly Leu Ser Ser Val Leu Ile Gly Cys Ala Met Asn
            20                  25                  30
Pro Ser Ala Glu Thr Lys Lys Pro Asn Asp Ala Lys Asn Gln Gln Pro
        35                  40                  45
Val Gln Thr His Glu Arg Met Thr Thr Ser Ser Glu His Val Thr Pro
    50                  55                  60
Leu Asp Phe Asn Tyr Pro Val His Ile Val Gln Ala Pro Gln Asn His
65                  70                  75                  80
His Val Val Gly Ile Leu Met Pro Arg Ile Gln Val Ser Asp Asn Leu
                85                  90                  95
Lys Pro Tyr Ile Asp Lys Phe Gln Asp Ala Leu Ile Asn Gln Ile Gln
            100                 105                 110
Thr Ile Phe Glu Lys Arg Gly Tyr Gln Val Leu Arg Phe Gln Asp Glu
        115                 120                 125
Lys Ala Leu Asn Val Gln Asp Lys Lys Lys Ile Phe Ser Val Leu Asp
    130                 135                 140
Leu Lys Gly Trp Val Gly Ile Leu Glu Asp Leu Lys Met Asn Leu Lys
145                 150                 155                 160
Asp Pro Asn Ser Pro Asn Leu Asp Thr Leu Val Asp Gln Ser Ser Gly
                165                 170                 175
Ser Val Trp Phe Asn Phe Tyr Glu Pro Glu Ser Asn Arg Val Val His
            180                 185                 190
Asp Phe Ala Val Glu Val Gly Thr Phe Gln Ala Ile Thr Tyr Thr Tyr
        195                 200                 205
Thr Ser Thr Asn Asn Ala Ser Gly Gly Phe Asn Ser Ser Lys Ser Val
    210                 215                 220
Ile His Glu Asn Leu Asp Lys Asn Arg Glu Asp Ala Ile His Lys Ile
225                 230                 235                 240
Leu Asn Arg Met Tyr Ala Val Val Met Lys Lys Ala Val Thr Glu Leu
                245                 250                 255
Thr Lys Glu Asn Ile Ala Lys Tyr Arg Asp Ala Ile Asp Arg Met Lys
            260                 265                 270
Gly Phe Lys Ser Ser Met Pro Gln Lys Lys
        275                 280
```

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 280 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...280

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

```
Met Lys Leu Arg Ala Ser Val Leu Ile Gly Val Ala Ile Leu Cys Leu
1               5                  10                  15
Ile Leu Ser Ala Cys Ser Asn Tyr Ala Lys Lys Val Val Lys Gln Lys
            20                  25                  30
Asn His Val Tyr Thr Pro Val Tyr Asn Glu Leu Ile Glu Lys Tyr Ser
        35                  40                  45
Glu Ile Pro Leu Asn Asp Lys Leu Lys Asp Thr Pro Phe Met Val Gln
    50                  55                  60
Val Lys Leu Pro Asn Tyr Lys Asp Tyr Leu Leu Asp Asn Lys Gln Val
65                  70                  75                  80
Val Leu Thr Phe Lys Leu Val His His Ser Lys Lys Ile Thr Leu Ile
                85                  90                  95
Gly Asp Ala Asn Lys Ile Leu Gln Tyr Lys Asn Tyr Phe Gln Ala Asn
            100                 105                 110
Gly Ala Arg Ser Asp Ile Asp Phe Tyr Leu Gln Pro Thr Leu Asn Gln
        115                 120                 125
Lys Gly Val Val Met Ile Ala Ser Asn Tyr Asn Asp Asn Pro Asn Asn
    130                 135                 140
Lys Glu Lys Pro Gln Thr Phe Asp Val Leu Gln Gly Ser Gln Pro Met
145                 150                 155                 160
Leu Gly Ala Asn Thr Lys Asn Leu His Gly Tyr Asp Val Ser Gly Ala
                165                 170                 175
Asn Asn Lys Gln Val Ile Asn Glu Val Ala Arg Glu Lys Ala Gln Leu
            180                 185                 190
Glu Lys Ile Asn Gln Tyr Tyr Lys Thr Leu Leu Gln Asp Lys Glu Gln
        195                 200                 205
Glu Tyr Thr Thr Arg Lys Asn Asn Gln Arg Glu Ile Leu Glu Thr Leu
    210                 215                 220
Ser Asn Arg Ala Gly Tyr Gln Met Arg Gln Asn Val Ile Ser Ser Glu
225                 230                 235                 240
Ile Phe Lys Asn Gly Asn Leu Asn Met Gln Ala Lys Glu Glu Glu Val
                245                 250                 255
Arg Glu Lys Leu Gln Glu Glu Arg Glu Asn Glu Tyr Leu Arg Asn Gln
            260                 265                 270
Ile Arg Ser Leu Leu Ser Gly Lys
        275                 280
```

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 393 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...393

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

```
Met Arg Lys Leu Phe Ile Pro Leu Leu Leu Phe Ser Ala Leu Glu Ala
1               5                  10                  15
Asn Glu Lys Asn Gly Phe Phe Ile Glu Ala Gly Phe Glu Thr Gly Leu
            20                  25                  30
Leu Glu Gly Thr Gln Thr Gln Glu Lys Arg His Thr Thr Thr Lys Asn
        35                  40                  45
Thr Tyr Ala Thr Tyr Asn Tyr Leu Pro Thr Asp Thr Ile Leu Lys Arg
    50                  55                  60
Ala Ala Asn Leu Phe Thr Asn Ala Glu Ala Ile Ser Lys Leu Lys Phe
65                  70                  75                  80
Ser Ser Leu Ser Pro Val Arg Val Leu Tyr Met Tyr Asn Gly Gln Leu
                85                  90                  95
Thr Ile Glu Asn Phe Leu Pro Tyr Asn Leu Asn Asn Val Lys Leu Ser
            100                 105                 110
Phe Thr Asp Ala Gln Gly Asn Val Ile Asp Leu Gly Val Ile Glu Thr
        115                 120                 125
Ile Pro Lys His Ser Lys Ile Val Leu Pro Gly Glu Ala Phe Asp Ser
    130                 135                 140
Leu Lys Ile Asp Pro Tyr Thr Leu Phe Leu Pro Lys Ile Glu Ala Thr
145                 150                 155                 160
Ser Thr Ser Ile Ser Asp Ala Asn Thr Gln Arg Val Phe Glu Thr Leu
                165                 170                 175
Asn Lys Ile Lys Thr Asn Leu Val Val Asn Tyr Arg Asn Glu Asn Lys
            180                 185                 190
Phe Lys Asp His Glu Asn His Trp Glu Ala Phe Thr Pro Gln Thr Ala
        195                 200                 205
Glu Glu Phe Thr Asn Leu Met Leu Asn Met Ile Ala Val Leu Asp Ser
    210                 215                 220
Gln Ser Trp Gly Asp Ala Ile Leu Asn Ala Pro Phe Glu Phe Thr Asn
225                 230                 235                 240
Ser Pro Thr Asp Cys Asp Asn Asp Pro Ser Lys Cys Val Asn Pro Gly
                245                 250                 255
Thr Asn Gly Leu Val Asn Ser Lys Val Asp Gln Lys Tyr Val Leu Asn
            260                 265                 270
```

Lys Gln Asp Ile Val Asn Lys Phe Lys Asn Lys Ala Asp Leu Asp Val
275 280 285
Ile Val Leu Lys Asp Ser Gly Val Val Gly Leu Gly Ser Asp Ile Thr
290 295 300
Pro Ser Asn Asn Asp Asp Gly Lys His Tyr Gly Gln Leu Gly Val Val
305 310 315 320
Ala Ser Ala Leu Asp Pro Lys Lys Leu Phe Gly Asp Asn Leu Lys Thr
325 330 335
Ile Asn Leu Glu Asp Leu Arg Thr Ile Leu His Glu Phe Ser His Thr
340 345 350
Lys Gly Tyr Gly His Asn Gly Asn Met Thr Tyr Gln Arg Val Pro Val
355 360 365
Thr Lys Asp Gly Gln Val Glu Lys Asp Ser Asn Gly Lys Pro Lys Asp
370 375 380
Ser Asp Gly Leu Pro Tyr Asn Val Cys
385 390

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 270 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...270

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Met Lys Lys Phe Val Ala Leu Gly Leu Leu Ser Ala Val Leu Ser Ser
1 5 10 15
Ser Leu Leu Ala Glu Gly Asp Gly Val Tyr Ile Gly Thr Asn Tyr Gln
20 25 30
Leu Gly Gln Ala Arg Leu Asn Ser Asn Ile Tyr Asn Thr Gly Asp Cys
35 40 45
Thr Gly Ser Val Val Gly Cys Pro Pro Gly Leu Thr Ala Asn Lys His
50 55 60
Asn Pro Gly Gly Thr Asn Ile Asn Trp His Ser Lys Tyr Ala Asn Gly
65 70 75 80
Ala Leu Asn Gly Phe Gly Leu Asn Val Gly Tyr Lys Lys Phe Phe Gln
85 90 95
Phe Lys Ser Leu Asp Met Thr Ser Lys Trp Phe Gly Phe Arg Val Tyr
100 105 110
Gly Leu Phe Asp Tyr Gly His Ala Asp Leu Gly Lys Gln Val Tyr Ala
115 120 125
Pro Asn Lys Ile Gln Leu Asp Met Val Ser Trp Gly Val Gly Ser Asp
130 135 140

```
Leu Leu Ala Asp Ile Ile Asp Lys Asp Asn Ala Ser Phe Gly Ile Phe
145                 150                 155                 160
Gly Gly Val Ala Ile Gly Gly Asn Thr Trp Lys Ser Ser Ala Ala Asn
                165                 170                 175
Tyr Trp Lys Glu Gln Ile Ile Glu Ala Lys Gly Pro Asp Val Cys Thr
            180                 185                 190
Pro Thr Tyr Cys Asn Pro Asn Ala Pro Tyr Ser Thr Asn Thr Ser Thr
        195                 200                 205
Val Ala Phe Gln Val Trp Leu Asn Phe Gly Val Arg Ala Asn Ile Tyr
    210                 215                 220
Lys His Asn Gly Val Glu Phe Gly Val Arg Val Pro Leu Leu Ile Asn
225                 230                 235                 240
Lys Phe Leu Ser Ala Gly Pro Asn Ala Thr Asn Leu Tyr Tyr His Leu
                245                 250                 255
Lys Arg Asp Tyr Ser Leu Tyr Leu Gly Tyr Asn Tyr Thr Phe
            260                 265                 270
```

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 140 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...140

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

```
Met His Pro Ile Met Phe Ala Tyr Ile Ala Asn Ala Leu Ala Gln Ala
1               5                   10                  15
Arg Lys Ile Asn Gly Thr Leu Cys Met Ala Phe Gln Lys Ile Ser Gln
            20                  25                  30
Val Lys Glu Leu Gly Ile Asp Lys Ala Lys Ser Leu Ile Gly Asn Leu
        35                  40                  45
Ser Gln Val Ile Ile Tyr Pro Thr Lys Asp Thr Asp Glu Leu Ile Glu
    50                  55                  60
Cys Gly Val Pro Leu Ser Asp Ser Glu Ile Asn Phe Leu His Asn Thr
65                  70                  75                  80
Asp Met Arg Ala Arg Gln Val Leu Val Lys Asn Ile Val Thr Asn Ala
                85                  90                  95
Ser Ala Phe Ile Glu Ile Asp Leu Lys Lys Ile Cys Lys Asn Tyr Phe
            100                 105                 110
Ile Phe Leu Ile Ala Met Leu Val Ile Glu Lys Ser Ser Met Ile Leu
        115                 120                 125
Lys Lys Gln Thr Lys Lys Leu Ile Arg Lys Ser Ile
    130                 135                 140
```

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 256 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...256

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

```
Met Leu Gly Ser Val Lys Lys Ala Val Phe Arg Val Leu Cys Leu Gly
1               5                  10                  15
Ala Leu Cys Leu Cys Gly Gly Leu Met Ala Glu Gln Asp Pro Lys Glu
            20                  25                  30
Leu Ile Phe Ser Gly Ile Thr Ile Tyr Thr Asp Lys Asn Phe Thr Arg
        35                  40                  45
Ala Lys Lys Tyr Phe Glu Lys Ala Cys Lys Ser Asn Asp Ala Asp Gly
    50                  55                  60
Cys Ala Ile Leu Arg Glu Val Tyr Ser Ser Gly Lys Ala Ile Ala Arg
65                  70                  75                  80
Glu Asn Ala Arg Glu Ser Ile Glu Lys Ala Leu Glu His Thr Ala Thr
                85                  90                  95
Ala Lys Val Cys Lys Leu Asn Asp Ala Glu Lys Cys Lys Asp Leu Ala
            100                 105                 110
Glu Phe Tyr Phe Asn Val Asn Asp Leu Lys Asn Ala Leu Glu Tyr Tyr
        115                 120                 125
Ser Lys Ser Cys Lys Leu Asn Asn Val Glu Gly Cys Met Leu Ser Ala
    130                 135                 140
Thr Phe Tyr Asn Asp Met Ile Lys Gly Leu Lys Lys Asp Lys Lys Asp
145                 150                 155                 160
Leu Glu Tyr Tyr Ser Lys Ala Cys Glu Leu Asn Asn Gly Gly Gly Cys
                165                 170                 175
Ser Lys Leu Gly Gly Asp Tyr Phe Phe Gly Glu Gly Val Thr Lys Asp
            180                 185                 190
Phe Lys Lys Ala Phe Glu Tyr Ser Ala Lys Ala Cys Glu Leu Asn Asp
        195                 200                 205
Ala Lys Gly Cys Tyr Ala Leu Ala Ala Phe Tyr Asn Glu Gly Lys Gly
    210                 215                 220
Val Ala Lys Asp Glu Lys Gln Thr Thr Glu Asn Leu Glu Lys Ser Cys
225                 230                 235                 240
Lys Leu Gly Leu Lys Glu Ala Cys Asp Ile Leu Lys Glu Gln Lys Gln
                245                 250                 255
```

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 242 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...242

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

```
Met Lys Lys Phe Phe Ser Gln Ser Leu Leu Ala Leu Ile Ile Ser Met
1               5                  10                  15
Asn Ala Val Ser Gly Met Asp Gly Asn Gly Val Phe Leu Gly Ala Gly
            20                  25                  30
Tyr Leu Gln Gly Gln Ala Gln Met His Ala Asp Ile Asn Ser Gln Lys
        35                  40                  45
Gln Ala Thr Asn Ala Thr Ile Lys Gly Phe Asp Ala Leu Leu Gly Tyr
    50                  55                  60
Gln Phe Phe Phe Glu Lys His Phe Gly Leu Arg Leu Tyr Gly Phe Phe
65                  70                  75                  80
Asp Tyr Ala His Ala Asn Ser Ile Lys Leu Lys Asn Pro Asn Tyr Asn
                85                  90                  95
Ser Glu Ala Ala Gln Val Ala Ser Gln Ile Leu Gly Lys Gln Glu Ile
            100                 105                 110
Asn Arg Leu Thr Asn Ile Ala Asp Pro Arg Thr Phe Glu Pro Asn Met
        115                 120                 125
Leu Thr Tyr Gly Gly Ala Met Asp Val Met Val Asn Val Ile Asn Asn
    130                 135                 140
Gly Ile Met Ser Leu Gly Ala Phe Gly Gly Ile Gln Leu Ala Gly Asn
145                 150                 155                 160
Ser Trp Leu Met Ala Thr Pro Ser Phe Glu Gly Ile Leu Val Glu Gln
                165                 170                 175
Ala Leu Val Ser Lys Lys Ala Thr Ser Phe Gln Phe Leu Phe Asn Val
            180                 185                 190
Gly Ala Arg Leu Arg Ile Leu Lys His Ser Ser Ile Glu Ala Gly Val
        195                 200                 205
Lys Phe Pro Met Leu Lys Lys Asn Pro Tyr Ile Thr Ala Lys Asn Leu
    210                 215                 220
Asp Ile Gly Phe Arg Arg Val Tyr Ser Trp Tyr Val Asn Tyr Val Phe
225                 230                 235                 240
Thr Phe
```

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 267 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...267

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

```
Met Asn Tyr Pro Asn Leu Pro Asn Ser Ala Leu Glu Ile Ser Glu Gln
1               5                  10                  15
Pro Glu Val Lys Glu Ile Thr Asn Glu Leu Leu Lys Gln Leu Gln Asn
            20                  25                  30
Ala Leu Arg Ser Asn Ala His Phe Ser Glu Gln Val Glu Leu Ser Leu
        35                  40                  45
Lys Cys Ile Val Arg Ile Leu Glu Val Leu Leu Ser Leu Asp Phe Phe
    50                  55                  60
Lys Asn Ala Asn Glu Ile Asp Ser Ser Leu Arg Asn Ser Ile Glu Trp
65                  70                  75                  80
Leu Thr Asn Ala Gly Glu Ser Leu Lys Leu Lys Met Lys Glu Tyr Glu
                85                  90                  95
Arg Phe Phe Ser Glu Phe Asn Thr Ser Met His Ala Asn Glu Gln Glu
            100                 105                 110
Val Thr Asn Thr Leu Asn Ala Asn Ala Glu Asn Ile Lys Ser Glu Ile
        115                 120                 125
Lys Lys Leu Glu Asn Gln Leu Ile Glu Thr Thr Thr Arg Leu Leu Thr
    130                 135                 140
Ser Tyr Gln Ile Phe Leu Asn Gln Ala Arg Asp Asn Ala Asn Asn Gln
145                 150                 155                 160
Ile Thr Lys Asn Lys Thr Gln Ser Leu Glu Ala Ile Thr Gln Ala Lys
                165                 170                 175
Asn Asn Ala Asn Asn Glu Ile Ser Asn Asn Gln Thr Gln Ala Ile Thr
            180                 185                 190
Asn Ile Thr Glu Ala Lys Thr Asn Ala Asn Asn Glu Ile Ser Asn Asn
        195                 200                 205
Gln Thr Gln Ala Ile Thr Asn Ile Asn Glu Ala Lys Glu Ser Ala Thr
    210                 215                 220
Thr Gln Ile Asn Ala Asn Lys Gln Glu Ala Ile Asn Asn Ile Thr Gln
225                 230                 235                 240
Glu Lys Thr Gln Ala Thr Ser Glu Ile Thr Glu Ala Lys Lys Thr Asp
                245                 250                 255
His Tyr Gln Asn Ile Asp Phe Phe Glu Phe Glu
            260                 265
```

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 544 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...544

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

```
Val Ile Glu Thr Ile Pro Lys His Ser Lys Ile Val Leu Pro Gly Glu
1               5                  10                  15
Ala Phe Asp Ser Leu Lys Glu Ala Phe Asp Lys Ile Asp Pro Tyr Thr
            20                  25                  30
Phe Phe Phe Pro Lys Phe Glu Ala Thr Ser Thr Ser Ile Ser Asp Thr
        35                  40                  45
Asn Thr Gln Arg Val Phe Glu Thr Leu Asn Asn Ile Lys Thr Asn Leu
    50                  55                  60
Ile Met Lys Tyr Ser Asn Glu Asn Pro Asn Asn Phe Asn Thr Cys Pro
65                  70                  75                  80
Tyr Asn Asn Asn Gly Asn Thr Lys Asn Asp Cys Trp Gln Asn Phe Thr
                85                  90                  95
Pro Gln Thr Ala Glu Glu Phe Thr Asn Leu Met Leu Asn Met Ile Ala
            100                 105                 110
Val Leu Asp Ser Gln Ser Trp Gly Asp Ala Ile Leu Asn Ala Pro Phe
        115                 120                 125
Glu Phe Thr Asn Ser Ser Thr Asp Cys Asp Ser Asp Pro Ser Lys Cys
    130                 135                 140
Val Asn Pro Gly Val Asn Gly Arg Val Asp Thr Lys Val Asp Gln Gln
145                 150                 155                 160
Tyr Ile Leu Asn Lys Gln Gly Ile Ile Asn Asn Phe Arg Lys Lys Ile
                165                 170                 175
Glu Ile Asp Ala Val Val Leu Lys Asn Ser Gly Val Val Gly Leu Ala
            180                 185                 190
Asn Gly Tyr Gly Asn Asp Gly Glu Tyr Gly Thr Leu Gly Val Glu Ala
        195                 200                 205
Tyr Ala Leu Asp Pro Lys Lys Leu Phe Gly Asn Asp Leu Lys Thr Ile
    210                 215                 220
Asn Leu Glu Asp Leu Arg Thr Ile Leu His Glu Phe Ser His Thr Lys
225                 230                 235                 240
Gly Tyr Gly His Asn Gly Asn Met Thr Tyr Gln Arg Val Pro Val Thr
                245                 250                 255
Lys Asp Gly Gln Val Glu Lys Asp Ser Asn Gly Lys Pro Lys Asp Ser
            260                 265                 270
Asp Gly Leu Pro Tyr Asn Val Cys Ser Leu Tyr Gly Gly Ser Asn Gln
        275                 280                 285
Pro Ala Phe Pro Ser Asn Tyr Pro Asn Ser Ile Tyr His Asn Cys Ala
    290                 295                 300
```

```
Asp Val Pro Ala Gly Phe Leu Gly Val Thr Ala Ala Val Trp Gln Gln
305                 310                 315                 320
Leu Ile Asn Gln Asn Ala Leu Pro Ile Asn Tyr Ala Asn Leu Gly Ser
                325                 330                 335
Gln Thr Asn Tyr Asn Leu Asn Ala Ser Leu Asn Thr Gln Asp Leu Ala
            340                 345                 350
Asn Ser Met Leu Ser Thr Ile Gln Lys Thr Phe Val Thr Ser Ser Val
        355                 360                 365
Thr Asn His His Phe Ser Asn Ala Ser Gln Ser Phe Arg Ser Pro Ile
    370                 375                 380
Leu Gly Val Asn Ala Lys Ile Gly Tyr Gln Asn Tyr Phe Asn Asp Phe
385                 390                 395                 400
Ile Gly Leu Ala Tyr Tyr Gly Ile Ile Lys Tyr Asn Tyr Ala Lys Ala
                405                 410                 415
Val Asn Gln Lys Val Gln Gln Leu Ser Tyr Gly Gly Gly Ile Asp Leu
            420                 425                 430
Leu Leu Asp Phe Ile Thr Thr Tyr Ser Asn Lys Asn Ser Pro Thr Gly
        435                 440                 445
Ile Gln Thr Lys Arg Asn Phe Ser Ser Ser Phe Gly Ile Phe Gly Gly
    450                 455                 460
Leu Arg Gly Leu Tyr Asn Ser Tyr Tyr Val Leu Asn Lys Val Lys Gly
465                 470                 475                 480
Ser Gly Asn Leu Asp Val Ala Thr Gly Leu Asn Tyr Arg Tyr Lys His
                485                 490                 495
Ser Lys Tyr Ser Val Gly Ile Ser Ile Pro Leu Ile Gln Arg Lys Ala
            500                 505                 510
Ser Val Val Ser Ser Gly Gly Asp Tyr Thr Asn Ser Phe Val Phe Asn
        515                 520                 525
Glu Gly Ala Ser His Phe Lys Val Phe Phe Asn Tyr Gly Gly Cys Phe
    530                 535                 540
```

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 356 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...356

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

```
Leu Met Lys Ser Ile Leu Leu Phe Met Ile Phe Val Val Cys Gln Leu
1               5                  10                  15
Glu Gly Lys Lys Phe Ser Gln Asp Asn Phe Lys Val Asp Tyr Asn Tyr
            20                  25                  30
```

```
Tyr Leu Arg Lys Gln Asp Leu His Ile Ile Lys Thr Gln Asn Asp Leu
        35                  40                  45
Ser Asn Ala Trp Tyr Leu Pro Pro Gln Lys Ala Pro Lys Glu His Ser
    50                  55                  60
Trp Val Asp Phe Ala Lys Lys Tyr Leu Asn Met Met Asp Tyr Leu Gly
65                  70                  75                  80
Thr Tyr Phe Leu Pro Phe Tyr His Ser Phe Thr Pro Ile Phe Gln Trp
                85                  90                  95
Tyr His Pro Asn Ile Asn Pro Tyr Gln Arg Asn Glu Phe Lys Phe Gln
            100                 105                 110
Ile Ser Phe Arg Val Pro Val Phe Arg His Ile Leu Trp Thr Lys Gly
        115                 120                 125
Thr Leu Tyr Leu Ala Tyr Thr Gln Thr Asn Trp Phe Gln Ile Tyr Asn
    130                 135                 140
Asp Pro Gln Ser Ala Pro Met Arg Met Ile Asn Phe Met Pro Glu Leu
145                 150                 155                 160
Ile Tyr Val Tyr Pro Ile Asn Phe Lys Pro Phe Gly Gly Lys Ile Gly
                165                 170                 175
Asn Phe Ser Glu Ile Trp Ile Gly Trp Gln His Ile Ser Asn Gly Val
            180                 185                 190
Gly Gly Ala Gln Cys Tyr Gln Pro Phe Asn Lys Glu Gly Asn Pro Glu
        195                 200                 205
Asn Gln Phe Pro Gly Gln Pro Val Ile Val Lys Asp Tyr Asn Gly Gln
    210                 215                 220
Lys Asp Val Arg Trp Gly Gly Cys Xaa Ser Val Xaa Xaa Gly Asn Xaa
225                 230                 235                 240
Leu Cys Phe Val Leu Val Trp Glu Lys Gly Gly Leu Lys Ile Met Val
                245                 250                 255
Ala Tyr Trp Pro Tyr Val Pro Tyr Asp Gln Ser Asn Pro Gln Leu Ile
            260                 265                 270
Asp Tyr Met Gly Tyr Gly Asn Ala Lys Ile Asp Tyr Arg Arg Gly Arg
        275                 280                 285
His His Phe Glu Leu Gln Leu Tyr Asp Ile Phe Thr Gln Tyr Trp Arg
    290                 295                 300
Tyr Asp Arg Trp His Gly Ala Phe Arg Leu Gly Tyr Thr Tyr Arg Ile
305                 310                 315                 320
Asn Pro Phe Val Gly Ile Tyr Ala Gln Trp Phe Asn Gly Tyr Gly Asp
                325                 330                 335
Gly Leu Tyr Glu Tyr Asp Val Phe Ser Asn Arg Ile Gly Val Gly Ile
            340                 345                 350
Arg Leu Asn Pro
        355
```

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 675 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...675

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

```
Leu Ser Lys Gly Leu Ser Ile Gly Asn Lys Ile Ile Leu Cys Val Ala
1               5                  10                  15
Leu Ile Val Ile Val Cys Val Ser Ile Leu Gly Val Ser Leu Asn Ser
            20                  25                  30
Arg Val Lys Glu Ile Leu Lys Glu Ser Ala Leu His Ser Met Gln Asp
        35                  40                  45
Ser Leu His Phe Lys Val Lys Glu Val Gln Ser Val Leu Glu Asn Thr
    50                  55                  60
Tyr Thr Ser Met Gly Ile Val Lys Glu Met Leu Pro Glu Asp Thr Lys
65                  70                  75                  80
Arg Glu Ile Lys Ile Gln Leu Leu Lys Asn Phe Ile Leu Ala Asn Ser
                85                  90                  95
His Val Ala Gly Val Ser Met Phe Phe Lys Asp Arg Glu Asp Leu Arg
            100                 105                 110
Leu Thr Leu Leu Arg Asp Asn Asp Thr Ile Lys Leu Met Glu Asn Pro
        115                 120                 125
Ser Leu Gly Ser Asn Pro Leu Ala Gln Lys Ala Met Lys Asn Lys Glu
    130                 135                 140
Ile Ser Lys Ser Leu Pro Tyr Tyr Arg Lys Met Pro Asn Gly Ala Glu
145                 150                 155                 160
Val Tyr Gly Val Asp Ile Leu Leu Pro Leu Phe Lys Glu Asn Thr Gln
                165                 170                 175
Glu Val Val Gly Val Leu Met Ile Phe Phe Ser Ile Asp Ser Phe Ser
            180                 185                 190
Asn Glu Ile Thr Lys Asn Arg Ser Asp Leu Phe Leu Ile Gly Val Lys
        195                 200                 205
Gly Lys Val Leu Leu Ser Ala Asn Lys Ser Leu Gln Asp Lys Ser Ile
    210                 215                 220
Thr Glu Ile Tyr Lys Ser Val Pro Lys Ala Thr Asn Glu Val Met Ala
225                 230                 235                 240
Ile Leu Glu Asn Gly Ser Lys Ala Thr Leu Glu Tyr Leu Asp Pro Phe
                245                 250                 255
Ser His Lys Glu Asn Phe Leu Ala Val Glu Thr Phe Lys Met Leu Gly
            260                 265                 270
Lys Thr Glu Ser Lys Asp Asn Leu Asn Trp Met Ile Ala Leu Ile Ile
        275                 280                 285
Glu Lys Asp Lys Val Tyr Glu Gln Val Gly Ser Val Arg Phe Val Val
    290                 295                 300
Val Ala Ala Ser Ala Ile Met Val Leu Ala Leu Ile Ile Ala Ile Thr
305                 310                 315                 320
Leu Leu Met Arg Ala Ile Val Ser Asn Arg Leu Glu Val Val Ser Ser
                325                 330                 335
Thr Leu Ser His Phe Phe Lys Leu Leu Asn Asn Gln Ala His Ser Ser
            340                 345                 350
Asp Ile Lys Leu Val Glu Ala Arg Ser Asn Asp Glu Leu Gly Arg Met
        355                 360                 365
Gln Thr Ala Ile Asn Lys Asn Ile Leu Gln Thr Gln Lys Thr Met Gln
```

```
    370                 375                 380
Glu Asp Arg Gln Ala Val Gln Asp Thr Ile Lys Val Val Ser Asp Val
385                 390                 395                 400
Lys Ala Gly Asn Phe Ala Val Arg Ile Thr Ala Glu Pro Ala Ser Pro
                405                 410                 415
Asp Leu Lys Glu Leu Arg Asp Ala Leu Asn Gly Ile Met Asp Tyr Leu
            420                 425                 430
Gln Glu Ser Val Gly Thr His Met Pro Ser Ile Phe Lys Ile Phe Glu
        435                 440                 445
Ser Tyr Ser Gly Leu Asp Phe Arg Gly Arg Ile Gln Asn Ala Ser Gly
    450                 455                 460
Arg Val Glu Leu Val Thr Asn Ala Leu Gly Gln Glu Ile Gln Lys Met
465                 470                 475                 480
Leu Glu Thr Ser Ser Asn Phe Ala Lys Asp Leu Ala Asn Asp Ser Ala
                485                 490                 495
Asn Leu Lys Glu Cys Val Gln Asn Leu Glu Lys Ala Ser Asn Ser Gln
            500                 505                 510
His Lys Ser Leu Met Glu Thr Ser Lys Thr Ile Glu Asn Ile Thr Thr
        515                 520                 525
Ser Ile Gln Gly Val Ser Ser Gln Ser Glu Ala Met Ile Glu Gln Gly
    530                 535                 540
Lys Asp Ile Lys Ser Ile Val Glu Ile Ile Arg Asp Ile Ala Asp Gln
545                 550                 555                 560
Thr Asn Leu Leu Ala Leu Asn Ala Ala Ile Glu Ala Ala Arg Ala Gly
                565                 570                 575
Glu His Gly Arg Gly Phe Ala Val Val Ala Asp Glu Val Arg Lys Leu
            580                 585                 590
Ala Glu Arg Thr Gln Lys Ser Leu Ser Glu Ile Glu Ala Asn Ile Asn
        595                 600                 605
Ile Leu Val Gln Ser Ile Ser Asp Thr Ser Glu Ser Ile Lys Asn Gln
    610                 615                 620
Val Lys Glu Val Glu Glu Ile Asn Ala Ser Ile Glu Ala Leu Arg Ser
625                 630                 635                 640
Val Thr Glu Gly Asn Leu Lys Ile Ala Ser Asp Ser Leu Glu Ile Ser
                645                 650                 655
Gln Glu Ile Asp Lys Val Ser Asn Asp Ile Leu Glu Asp Val Asn Lys
            660                 665                 670
Lys Gln Phe
        675
```

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION 1...271

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

```
Met Asn Ile Phe Lys Arg Ile Ile Cys Val Thr Ala Ile Val Leu Gly
1               5                  10                  15
Phe Phe Asn Leu Leu Asp Ala Lys His His Lys Glu Lys Lys Glu Asp
            20                  25                  30
His Lys Ile Thr Arg Glu Leu Lys Val Gly Ala Asn Pro Val Pro His
        35                  40                  45
Ala Gln Ile Leu Gln Ser Val Val Asp Asp Leu Lys Glu Lys Gly Ile
    50                  55                  60
Lys Leu Val Ile Val Ser Phe Thr Asp Tyr Val Leu Pro Asn Leu Ala
65                  70                  75                  80
Leu Asn Asp Gly Ser Leu Asp Ala Asn Tyr Phe Gln His Arg Pro Tyr
                85                  90                  95
Leu Asp Arg Phe Asn Leu Asp Arg Lys Met His Leu Val Gly Leu Ala
            100                 105                 110
Asn Ile His Val Glu Pro Leu Arg Phe Tyr Ser Gln Lys Ile Thr Asp
        115                 120                 125
Ile Lys Asn Leu Lys Lys Gly Ser Val Ile Ala Val Pro Asn Asp Pro
    130                 135                 140
Ala Asn Gln Gly Arg Ala Leu Ile Leu Leu His Lys Gln Gly Leu Ile
145                 150                 155                 160
Ala Leu Lys Asp Pro Ser Asn Leu Tyr Ala Thr Glu Phe Asp Ile Val
                165                 170                 175
Lys Asn Pro Tyr Asn Ile Lys Ile Lys Pro Leu Glu Ala Ala Leu Leu
            180                 185                 190
Pro Lys Val Leu Gly Asp Val Asp Gly Ala Ile Ile Thr Gly Asn Tyr
        195                 200                 205
Ala Leu Gln Ala Lys Leu Thr Gly Ala Leu Phe Ser Glu Asp Lys Asp
    210                 215                 220
Ser Pro Tyr Ala Asn Leu Val Ala Ser Arg Glu Asp Asn Ala Gln Asp
225                 230                 235                 240
Glu Ala Ile Lys Ala Leu Ile Glu Ala Leu Gln Ser Glu Lys Thr Arg
                245                 250                 255
Lys Phe Ile Leu Asp Thr Tyr Lys Gly Ala Ile Ile Pro Ala Phe
            260                 265                 270
```

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 161 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION 1...161

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

```
Met Phe Phe Lys Thr Tyr Gln Lys Leu Leu Gly Ala Ser Cys Leu Ala
1               5                  10                  15
Leu Tyr Leu Val Gly Cys Gly Asn Gly Gly Gly Gly Glu Ser Pro Val
            20                  25                  30
Glu Met Ile Ala Asn Ser Glu Gly Thr Phe Gln Ile Asp Ser Lys Ala
        35                  40                  45
Asp Ser Ile Thr Ile Gln Gly Val Lys Leu Asn Arg Gly Asn Cys Ala
    50                  55                  60
Val Asn Phe Val Pro Val Ser Glu Thr Phe Gln Met Gly Val Leu Ser
65                  70                  75                  80
Gln Val Thr Pro Ile Ser Ile Gln Asp Phe Lys Asp Met Ala Ser Thr
                85                  90                  95
Tyr Lys Ile Phe Asp Gln Lys Lys Gly Leu Ala Asn Ile Ala Asn Lys
            100                 105                 110
Ile Ser Gln Leu Glu Gln Lys Gly Val Met Met Glu Pro Gln Thr Leu
        115                 120                 125
Asn Phe Gly Glu Ser Leu Lys Gly Ile Ser Gln Gly Cys Asn Ile Ile
    130                 135                 140
Glu Ala Glu Ile Gln Thr Asp Lys Gly Ala Trp Thr Phe Asn Phe Asp
145                 150                 155                 160
Lys
```

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 337 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...337

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

```
Met Ile Arg Leu Lys Gly Leu Asn Lys Thr Leu Lys Thr Ser Leu Leu
1               5                  10                  15
Ala Gly Val Leu Leu Gly Ala Thr Ala Pro Leu Met Ala Lys Pro Leu
            20                  25                  30
Leu Ser Asp Glu Asp Leu Leu Lys Arg Val Lys Leu His Asn Ile Lys
        35                  40                  45
Glu Asp Thr Leu Thr Ser Cys Asn Ala Lys Val Asp Gly Ser Gln Tyr
```

```
    50                  55                  60
Leu Asn Ser Gly Trp Asn Leu Ser Lys Glu Phe Pro Gln Glu Tyr Arg
65                  70                  75                  80
Glu Lys Ile Phe Glu Cys Val Glu Glu Glu Lys His Lys Gln Ala Leu
                85                  90                  95
Asn Leu Ile Asn Lys Glu Asp Thr Glu Asp Lys Glu Glu Leu Ala Lys
            100                 105                 110
Lys Ile Lys Glu Ile Lys Glu Lys Ala Lys Val Leu Arg Gln Lys Phe
        115                 120                 125
Met Ala Phe Glu Met Lys Glu His Ser Lys Glu Phe Pro Asn Lys Lys
    130                 135                 140
Gln Leu Gln Thr Met Leu Glu Asn Ala Phe Asp Asn Gly Ala Glu Ser
145                 150                 155                 160
Phe Ile Asp Asp Trp His Glu Arg Phe Gly Gly Ile Ser Arg Glu Asn
                165                 170                 175
Thr Tyr Lys Ala Leu Gly Ile Lys Glu Tyr Ser Asp Glu Gly Lys Ile
            180                 185                 190
Leu Ala Phe Gly Glu Arg Ser Tyr Ile Arg Gln Tyr Lys Lys Asp Phe
        195                 200                 205
Glu Glu Ser Thr Tyr Asp Thr Arg Gln Thr Leu Ser Ala Met Ala Asn
    210                 215                 220
Met Ser Gly Glu Asn Asp Tyr Lys Ile Thr Trp Leu Lys Pro Lys Tyr
225                 230                 235                 240
Gln Leu His Ser Ser Asn Asn Ile Lys Pro Leu Met Ser Asn Thr Glu
                245                 250                 255
Leu Leu Asn Met Ile Glu Leu Thr Asn Ile Lys Lys Glu Tyr Val Met
            260                 265                 270
Gly Cys Asn Met Glu Ile Asp Gly Ser Lys Tyr Pro Ile His Lys Asp
        275                 280                 285
Trp Gly Phe Phe Gly Lys Ala Lys Val Pro Glu Thr Trp Arg Asn Lys
    290                 295                 300
Ile Trp Glu Cys Ile Lys Asn Lys Val Lys Ser Tyr Asp Asn Thr Thr
305                 310                 315                 320
Ala Glu Ile Gly Ile Val Trp Lys Lys Asn Thr Tyr Ser Ile Ser His
                325                 330                 335
His
```

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 416 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...416

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

```
Met Lys Lys Leu Val Phe Ser Met Leu Leu Cys Cys Lys Ser Val Phe
1               5                  10                  15
Ala Glu Gly Glu Thr Pro Leu Ile Val Asn Asp Pro Glu Thr His Val
            20                  25                  30
Ser Gln Ala Thr Ile Ile Gly Lys Met Val Asp Ser Ile Lys Arg Tyr
        35                  40                  45
Glu Glu Ile Ile Ser Lys Ala Gln Ala Gln Val Asn Gln Leu Gln Lys
    50                  55                  60
Val Asn Asn Met Ile Asn Thr Thr Asn Ser Leu Ile Ser Ser Ser Ala
65                  70                  75                  80
Ile Thr Leu Ala Asn Pro Met Gln Val Leu Gln Asn Ala Gln Tyr Gln
                85                  90                  95
Ile Glu Ser Ile Arg Tyr Asn Tyr Glu Asn Leu Lys Gln Ser Ile Glu
            100                 105                 110
Asn Trp Asn Ala Gln Asn Leu Leu Arg Asn Lys Tyr Leu Gln Gln Gln
        115                 120                 125
Cys Pro Trp Leu Asn Val Asn Ala Leu Thr Asn Asn Lys Ile Val Asn
    130                 135                 140
Leu Lys Asp Leu Asn Asn Leu Ile Thr Lys Asn Gly Glu Gln Thr Gln
145                 150                 155                 160
Thr Ala Arg Asp Val Gln Asn Leu Ile Gln Ser Ile Ser Gly Ser Gly
                165                 170                 175
Tyr Gly Asn Met Gln Ser Leu Ala Gly Glu Leu Ser Gly Arg Ala Trp
            180                 185                 190
Gly Glu Met Leu Cys Lys Met Val Asn Asp Ser Asn Tyr Glu Ser Glu
        195                 200                 205
Gln Ala Leu Leu Ala Thr Gly Asn Asn Pro Glu Glu Gln Lys Arg Arg
    210                 215                 220
Phe Leu Leu Arg Val Lys Lys Lys Val Asn Asp Asn Lys Gln Leu Lys
225                 230                 235                 240
Asp Lys Leu Asp Pro Phe Leu Lys Arg Leu Asp Val Leu Gln Thr Glu
                245                 250                 255
Phe Gly Val Thr Asp Pro Thr Ala Asn His Asn Lys Gln Gly Ile His
            260                 265                 270
Tyr Cys Thr Glu Asn Lys Glu Thr Gly Lys Cys Asp Pro Ile Lys Asn
        275                 280                 285
Val Phe Arg Thr Thr Arg Leu Asp Asn Glu Leu Glu Gln Glu Ile Gln
    290                 295                 300
Thr Leu Thr Leu Asp Leu Ile Lys Ala Ser Asn Lys Asp Ala Gln Ser
305                 310                 315                 320
Gln Ala Tyr Ala Asn Phe Asn Gln Arg Ile Lys Leu Leu Thr Leu Lys
                325                 330                 335
Tyr Leu Lys Glu Ile Thr Asn Gln Met Leu Phe Leu Asn Gln Thr Met
            340                 345                 350
Ala Met Gln Ser Glu Ile Met Thr Asp Asp Tyr Phe Arg Gln Asn Asn
        355                 360                 365
Asp Gly Phe Gly Glu Lys Glu Asn His Ile Asp Lys Gln Leu Thr Gln
    370                 375                 380
Lys Arg Ile Asn Glu Arg Glu Arg Ala Arg Ile Tyr Phe Gln Asn Pro
385                 390                 395                 400
Asn Val Lys Phe Asp Gln Phe Gly Phe Pro Ile Phe Ser Ile Trp Asp
                405                 410                 415
```

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 376 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...376

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

```
Val Asn Lys Trp Ile Lys Gly Ala Val Val Phe Val Gly Gly Phe Ala
1               5                  10                  15
Thr Ile Thr Thr Phe Ser Leu Ile Tyr His Gln Lys Pro Lys Ala Pro
            20                  25                  30
Leu Asn Asn Gln Pro Ser Leu Leu Asn Asp Asp Glu Val Lys Tyr Pro
        35                  40                  45
Leu Gln Asp Tyr Thr Phe Thr Gln Asn Pro Gln Pro Thr Asn Thr Glu
    50                  55                  60
Ser Ser Lys Asp Ala Thr Ile Lys Ala Leu Gln Glu Gln Leu Lys Ala
65                  70                  75                  80
Ala Leu Lys Ala Leu Asn Ser Lys Glu Met Asn Tyr Ser Lys Glu Glu
                85                  90                  95
Thr Phe Thr Ser Pro Pro Met Asp Pro Lys Thr Thr Pro Pro Lys Lys
            100                 105                 110
Asp Phe Ser Pro Lys Gln Leu Asp Leu Leu Ala Ser Arg Ile Thr Pro
        115                 120                 125
Phe Lys Gln Ser Pro Lys Asn Tyr Glu Glu Asn Leu Ile Phe Pro Val
    130                 135                 140
Asp Asn Pro Asn Gly Ile Asp Ser Phe Thr Asn Leu Lys Glu Lys Asp
145                 150                 155                 160
Ile Ala Thr Asn Glu Asn Lys Leu Leu Arg Thr Ile Thr Ala Asp Lys
                165                 170                 175
Met Ile Pro Ala Phe Leu Ile Thr Pro Ile Ser Ser Gln Ile Ala Gly
            180                 185                 190
Lys Val Ile Ala Gln Val Glu Ser Asp Ile Phe Ala Ser Met Gly Lys
        195                 200                 205
Ala Val Leu Ile Pro Lys Gly Ser Lys Val Ile Gly Tyr Tyr Ser Asn
    210                 215                 220
Asn Asn Lys Met Gly Glu Tyr Arg Leu Asp Ile Val Trp Ser Arg Ile
225                 230                 235                 240
Ile Thr Pro His Gly Ile Asn Ile Met Leu Thr Asn Ala Lys Gly Ala
                245                 250                 255
Asp Ile Lys Gly Tyr Asn Gly Leu Val Gly Glu Leu Ile Glu Arg Asn
            260                 265                 270
```

```
Phe Gln Arg Tyr Gly Val Pro Leu Leu Leu Ser Thr Leu Thr Asn Gly
        275                 280                 285
Leu Leu Ile Gly Ile Thr Ser Ala Leu Asn Asn Arg Gly Asn Lys Glu
    290                 295                 300
Glu Val Thr Asn Phe Phe Gly Asp Tyr Leu Leu Leu Gln Leu Met Arg
305                 310                 315                 320
Gln Ser Gly Met Gly Ile Asn Gln Val Val Asn Gln Ile Leu Arg Asp
                325                 330                 335
Lys Ser Lys Ile Ala Pro Ile Val Val Ile Arg Glu Gly Ser Arg Val
            340                 345                 350
Phe Ile Ser Pro Asn Thr Asp Ile Phe Phe Pro Ile Pro Arg Glu Asn
        355                 360                 365
Glu Val Ile Ala Glu Phe Leu Lys
    370                 375
```

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 916 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...916

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

```
Val Asp Leu Arg Ile Gln Ser Lys Glu Val Ser His Asn Leu Lys Glu
1               5                   10                  15
Leu Ser Lys Thr Leu Ile Ser Tyr Pro Phe Glu Lys His Val Glu Ala
            20                  25                  30
Leu Gly Glu Gln Cys Ser Asn Phe Val Ser Ile Pro Ile Asn Asn Asp
        35                  40                  45
Asp Tyr Ser Asn Ile Cys Thr Phe Val Ser Asp Phe Ile Asn Leu Ile
    50                  55                  60
Ala Ser Tyr Asn Leu Leu Glu Ser Phe Leu Asp Phe Tyr Lys Asp Lys
65                  70                  75                  80
Leu Lys Leu Ser Glu Leu Val Thr Glu Tyr Ala Asn Val Thr Asn Asn
                85                  90                  95
Leu Leu Phe Lys Lys Leu Ile Lys His Leu Ser Gly Asn Asn Gln Leu
            100                 105                 110
Val Lys Asn Phe Tyr Gln Cys Ile Arg Glu Ile Ile Lys Tyr Asn Ala
        115                 120                 125
Pro Asn Lys Glu Tyr Lys Pro Asn Gln Phe Phe Ile Ile Gly Lys Gly
    130                 135                 140
Lys Gln Lys Gln Leu Ala Lys Ile Tyr Ser His Leu Lys Glu Leu Ser
145                 150                 155                 160
```

Ala Ser Glu Ile Lys Pro Gln Asp Met Glu Asp Ile Leu Lys Lys Leu
165 170 175
Glu Glu Leu Asp Lys Ile Phe Lys Thr Thr Asp Phe Thr Lys Phe Thr
180 185 190
Pro Lys Thr Glu Ile Lys Asp Ile Ile Lys Glu Ile Asp Glu Lys Tyr
195 200 205
Pro Ile Asn Glu Asn Phe Lys Arg Gln Phe Asn Glu Phe Glu Ser Asn
210 215 220
Ile Glu Lys His Asp Glu Ile Lys Lys Asp Phe Glu Arg Asn Lys Glu
225 230 235 240
Ser Leu Ile Arg Glu Ile Glu Asn His Cys Lys Asn Glu Cys Asn Ser
245 250 255
Glu Glu Glu Pro Glu Tyr Lys Ile Asn Asp Leu Leu Lys Asn Ile Gln
260 265 270
Gln Ile Cys Lys Asn Tyr Ile Glu Ser His Ala Val Asn Asp Val Ser
275 280 285
Lys Asp Ile Lys Ser Met Met Cys Gln Phe Tyr Leu Lys Gln Ile Asp
290 295 300
Leu Leu Val Asn Ser Glu Ile Val Arg Tyr Arg Tyr Ser Asn Leu Phe
305 310 315 320
Glu Pro Ile Gln Arg Ser Leu Trp Glu Ser Ile Lys Ile Leu Asp Asn
325 330 335
Glu Ser Gly Ile Tyr Leu Phe Pro Lys Asn Ile Gly Glu Ile Lys Asp
340 345 350
Lys Phe Glu Ala Asn Lys Glu Lys Phe Lys Gln Ser Lys Asn Val Ser
355 360 365
Glu Phe Ala Glu Tyr Cys Arg Glu Cys Asn Pro Tyr Thr Ala Phe Asn
370 375 380
Phe His Leu Asn Ile Asn Asn Gly Leu Ser His Gln Phe Glu Lys Phe
385 390 395 400
Val Pro Ile Met Lys Glu Tyr Lys Glu Pro Lys Ile Thr Asp Asn Asp
405 410 415
Leu Glu Ala Ile Ser Thr Lys Glu Thr Gly Leu Ala Ser Gln Leu Ser
420 425 430
Gly His Trp Phe Phe Gln Leu Ser Leu Phe Asn Lys Thr Asn Phe Asn
435 440 445
Pro Asn Lys Ile Trp Ile Pro Leu Glu Phe Asn Lys Arg Ser Lys Ile
450 455 460
Lys Phe Asp Lys Asp Leu Glu Ile Tyr Phe Asp Ser His Glu Ser Phe
465 470 475 480
Asn Ile Ser Lys Lys Tyr Leu Gln Glu Ile Asp Gln Glu Ser Leu Lys
485 490 495
Lys Ile Lys Gln Ser Lys Asp Phe Phe Ser Ile Gln Lys Ile Glu Ser
500 505 510
Lys His Asp Asn Asn Asp Ile Leu Gln Leu Glu Phe Phe Glu Asn Asp
515 520 525
Thr Ser Phe Leu Phe Ala Lys Gly Ser Phe Ala Glu Ile Leu Glu Tyr
530 535 540
Asn Met Gln Leu Lys Ile Asp Ser Leu Ile Thr Lys Glu Phe Asn Lys
545 550 555 560
Leu Leu Ala Ile Val Gln Asp Ser Pro Gln Asp Ser Tyr Gln Leu Lys
565 570 575
Ile Arg Val Arg His Asn Asn Lys Leu Pro Arg Glu Lys Tyr Thr Glu
580 585 590
His Glu Ile Lys Leu Glu Val Tyr Asp Cys Arg Lys Ser His Asp His

595 600 605
Asn Glu Pro Ile Ile Leu Ser Gln Gln Ser Thr Gly Phe Gln Trp Ala
610 615 620
Phe Asn Phe Met Phe Gly Phe Leu Tyr Asn Val Gly Ser His Phe Ser
625 630 635 640
Phe Asn His Asn Ile Ile Tyr Val Met Asp Glu Pro Ala Thr His Leu
645 650 655
Ser Val Pro Ala Arg Lys Glu Phe Arg Lys Phe Leu Lys Glu Tyr Ala
660 665 670
His Lys Asn His Val Thr Phe Val Leu Ala Thr His Asp Pro Phe Leu
675 680 685
Val Asp Thr Asp His Leu Asp Glu Ile Arg Ile Val Glu Lys Glu Thr
690 695 700
Glu Gly Ser Val Ile Lys Asn His Phe Asn Tyr Pro Leu Asn Asn Ala
705 710 715 720
Ser Lys Asp Ser Asp Ala Leu Asp Lys Ile Lys Arg Ser Leu Gly Val
725 730 735
Gly Gln His Val Phe His Asn Pro Gln Lys His Arg Ile Ile Phe Val
740 745 750
Glu Gly Ile Thr Asp Tyr Cys Tyr Leu Ser Ala Phe Lys Leu Tyr Leu
755 760 765
Arg Tyr Lys Glu Tyr Lys Asp Asn Pro Ile Pro Phe Thr Phe Leu Pro
770 775 780
Ile Ser Gly Leu Lys Asn Asp Ser Asn Asp Met Lys Glu Thr Ile Glu
785 790 795 800
Lys Leu Cys Glu Leu Asp Asn His Pro Ile Val Leu Thr Asp Asp Asp
805 810 815
Arg Lys Cys Val Phe Asn Gln Gln Ala Thr Ser Glu Arg Phe Lys Arg
820 825 830
Ala Asn Glu Glu Met His Asp Pro Ile Thr Ile Leu Gln Leu Ser Asp
835 840 845
Cys Asp Arg His Phe Lys Gln Ile Glu Asp Cys Phe Ser Ala Asn Asp
850 855 860
Arg Asn Lys Tyr Ala Lys Asn Lys Gln Met Glu Leu Ser Met Ala Phe
865 870 875 880
Lys Thr Arg Leu Leu Tyr Gly Gly Glu Asp Ala Ile Glu Lys Gln Thr
885 890 895
Lys Arg Asn Phe Leu Lys Leu Phe Lys Trp Ile Ala Trp Ala Thr Asn
900 905 910
Leu Ile Lys Asn
915

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 176 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...176

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

```
Met Thr Ala Met Met Arg Tyr Phe His Ile Tyr Ala Thr Thr Phe Phe
1               5                  10                  15
Phe Pro Leu Ala Leu Leu Phe Ala Val Ser Gly Leu Ser Leu Leu Phe
            20                  25                  30
Lys Ala Arg Gln Asp Thr Gly Ala Lys Ile Lys Glu Trp Val Leu Glu
        35                  40                  45
Lys Ser Leu Lys Lys Glu Glu Arg Leu Asp Phe Leu Lys Gly Phe Ile
    50                  55                  60
Lys Glu Asn His Ile Ala Met Pro Lys Lys Ile Glu Pro Arg Glu Tyr
65                  70                  75                  80
Arg Gly Ala Leu Val Ile Gly Thr Pro Leu Tyr Glu Ile Asn Leu Glu
                85                  90                  95
Thr Lys Gly Thr Gln Thr Lys Ile Lys Thr Ile Glu Arg Gly Phe Leu
            100                 105                 110
Gly Ala Leu Ile Met Leu His Lys Ala Lys Val Gly Ile Val Phe Gln
        115                 120                 125
Ala Leu Leu Gly Ile Phe Cys Val Phe Leu Leu Leu Phe Tyr Leu Ser
    130                 135                 140
Ala Phe Leu Met Val Ala Phe Lys Asp Thr Lys Arg Met Phe Ile Ser
145                 150                 155                 160
Val Leu Ile Gly Ser Val Val Phe Phe Gly Ala Ile Tyr Trp Ser Leu
                165                 170                 175
```

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 222 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...222

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

```
Met Phe Lys Asn Ala Leu Asn Ile Gln Asp Phe Ser Phe Lys Asn His
1               5                  10                  15
Thr Ser Thr Ala Ile Ile Gly Thr Asn Gly Ala Gly Lys Ser Thr Leu
            20                  25                  30
Ile Asn Thr Ile Leu Gly Ile Arg Ser Asp Tyr Asn Phe Lys Ala Gln
```

```
          35                  40                  45
Asn Asn Asn Ile Pro Tyr His Asp Asn Val Ile Pro Gln Arg Lys Gln
    50                  55                  60
Leu Gly Val Val Ser Asn Leu Phe Asn Tyr Pro Pro Gly Leu Asn Ala
65                  70                  75                  80
Asn Asp Leu Phe Lys Phe Tyr Gln Phe Phe His Lys Asn Cys Thr Leu
                85                  90                  95
Asp Leu Phe Glu Lys Asn Leu Leu Asn Lys Thr Tyr Glu His Leu Ser
            100                 105                 110
Asp Gly Gln Lys Gln Arg Leu Lys Ile Asp Leu Ala Leu Ser His His
        115                 120                 125
Pro Gln Leu Val Ile Met Asp Glu Pro Glu Thr Ser Leu Glu Gln Asn
    130                 135                 140
Ala Leu Ile Arg Leu Ser Asn Leu Ile Ser Leu Arg Asn Thr Gln Gln
145                 150                 155                 160
Leu Thr Ser Ile Ile Ala Thr His Asp Pro Ile Val Leu Asp Ser Cys
                165                 170                 175
Glu Trp Val Leu Leu Leu Lys Asn Gly Asn Ile Ala Gln Tyr Lys Pro
            180                 185                 190
Leu Asn Ser Ile Leu Lys Ser Val Ala Lys Thr Phe Asn Phe Lys Glu
        195                 200                 205
Lys Pro Thr Thr Lys Asp Leu Leu Ala Leu Leu Lys Asp Ile
    210                 215                 220
```

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 406 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...406

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

```
Met Tyr Ala Ala His Pro Ile Lys Pro Ile Lys Ala Pro Lys Leu Lys
1               5                   10                  15
Ser Gln Phe Leu Arg Arg Val Phe Val Gly Ala Ser Ile Arg Arg Trp
            20                  25                  30
Asn Asp Gln Ala Cys Pro Leu Glu Phe Val Glu Leu Asp Lys Gln Ala
        35                  40                  45
His Lys Ala Met Ile Ala Tyr Leu Leu Ala Lys Asp Leu Lys Asp Arg
    50                  55                  60
Gly Lys Asp Leu Asp Leu Asp Leu Leu Ile Lys Tyr Phe Cys Phe Glu
65                  70                  75                  80
Phe Leu Glu Arg Leu Val Leu Thr Asp Ile Lys Pro Pro Ile Phe Tyr
```

```
                    85                  90                  95
Ala Leu Gln Gln Thr His Ser Lys Glu Leu Ala Ser Tyr Val Ala Gln
            100                 105                 110
Ser Leu Gln Asp Glu Ile Ser Ala Tyr Phe Ser Leu Glu Glu Leu Lys
        115                 120                 125
Glu Tyr Leu Ser His Arg Pro Gln Ile Leu Glu Thr Gln Ile Leu Glu
    130                 135                 140
Ser Ala His Phe Tyr Ala Ser Lys Trp Glu Phe Asp Ile Ile Tyr His
145                 150                 155                 160
Phe Asn Pro Asn Met Tyr Gly Val Lys Glu Ile Lys Asp Lys Ile Asp
                165                 170                 175
Lys Gln Leu His Asn Asn Asp His Leu Phe Glu Gly Leu Phe Gly Glu
            180                 185                 190
Lys Glu Asp Leu Lys Lys Leu Val Ser Met Phe Gly Gln Leu Arg Phe
        195                 200                 205
Gln Lys Arg Trp Ser Gln Thr Pro Arg Val Pro Gln Thr Ser Val Leu
    210                 215                 220
Gly His Thr Leu Cys Val Ala Ile Met Gly Tyr Leu Leu Ser Phe Asp
225                 230                 235                 240
Leu Lys Ala Cys Lys Ser Met Arg Ile Asn His Phe Leu Gly Gly Leu
                245                 250                 255
Phe His Asp Leu Pro Glu Ile Leu Thr Arg Asp Ile Ile Thr Pro Ile
            260                 265                 270
Lys Gln Ser Val Ala Gly Leu Asp His Cys Ile Lys Glu Ile Glu Lys
        275                 280                 285
Lys Glu Met Gln Asn Lys Val Tyr Ser Phe Val Ser Leu Gly Val Gln
    290                 295                 300
Glu Asp Leu Lys Tyr Phe Thr Glu Asn Glu Phe Lys Asn Arg Tyr Lys
305                 310                 315                 320
Asp Lys Ser His Gln Ile Val Phe Thr Lys Asp Ala Glu Glu Leu Phe
                325                 330                 335
Thr Leu Tyr Asn Ser Asp Glu Tyr Leu Gly Val Cys Gly Glu Leu Leu
            340                 345                 350
Lys Val Cys Asp His Leu Ser Ala Phe Leu Glu Ala Gln Ile Ser Leu
        355                 360                 365
Ser His Gly Ile Ser Ser Tyr Asp Leu Ile Gln Gly Ala Lys Asn Leu
    370                 375                 380
Leu Glu Leu Arg Ser Gln Thr Glu Leu Leu Asp Leu Asp Leu Gly Lys
385                 390                 395                 400
Leu Phe Arg Asp Phe Lys
                405
```

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 335 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...335

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

```
Val Leu Trp Val Leu Tyr Phe Leu Thr Ser Leu Phe Ile Cys Ser Leu
1               5                  10                  15
Ile Val Leu Trp Ser Lys Lys Ser Met Leu Phe Val Asp Asn Ala Asn
            20                  25                  30
Lys Ile Gln Gly Phe His His Ala Arg Thr Pro Arg Ala Gly Gly Leu
        35                  40                  45
Gly Ile Phe Leu Ser Phe Ala Leu Ala Cys Tyr Leu Glu Pro Phe Glu
    50                  55                  60
Met Pro Phe Lys Gly Pro Phe Val Phe Leu Gly Leu Ser Leu Val Phe
65                  70                  75                  80
Leu Ser Gly Phe Leu Glu Asp Ile Asn Leu Ser Leu Ser Pro Lys Ile
                85                  90                  95
Arg Leu Ile Leu Gln Ala Val Gly Val Val Cys Ile Ile Ser Ser Thr
            100                 105                 110
Pro Leu Val Val Ser Asp Phe Ser Pro Leu Phe Ser Leu Pro Tyr Phe
        115                 120                 125
Ile Ala Phe Leu Phe Ala Ile Phe Met Leu Val Gly Ile Ser Asn Ala
    130                 135                 140
Ile Asn Ile Ile Asp Gly Phe Asn Gly Leu Ala Ser Gly Ile Cys Ala
145                 150                 155                 160
Ile Ala Leu Leu Val Ile His Tyr Ile Asp Pro Ser Ser Leu Ser Cys
                165                 170                 175
Leu Leu Ala Tyr Met Val Leu Gly Phe Met Val Leu Asn Phe Pro Ser
            180                 185                 190
Gly Lys Ile Phe Leu Gly Asp Gly Gly Ala Tyr Phe Leu Gly Leu Val
        195                 200                 205
Cys Gly Ile Ser Leu Leu His Leu Ser Leu Glu Gln Lys Ile Ser Val
    210                 215                 220
Phe Phe Gly Leu Asn Leu Met Leu Tyr Pro Val Ile Glu Val Leu Phe
225                 230                 235                 240
Ser Ile Leu Arg Arg Lys Ile Lys Arg Gln Lys Ala Thr Met Pro Asp
                245                 250                 255
Asn Leu His Leu His Thr Leu Leu Phe Lys Phe Leu Gln Gln Arg Ser
            260                 265                 270
Phe Asn Tyr Pro Asn Pro Leu Cys Ala Phe Ile Leu Ile Leu Cys Asn
        275                 280                 285
Leu Pro Phe Ile Leu Ile Ser Val Leu Phe Arg Leu Asp Ala Tyr Ala
    290                 295                 300
Leu Ile Val Ile Ser Leu Val Phe Ile Ala Cys Tyr Leu Ile Gly Tyr
305                 310                 315                 320
Ala Tyr Leu Asn Arg Gln Val Cys Ala Leu Glu Lys Arg Ala Phe
                325                 330                 335
```

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 amino acids
(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...96

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

```
Met Lys Lys Val Ile Val Ala Leu Gly Val Leu Ala Phe Ala Asn Val
1               5                  10                  15
Leu Met Ala Thr Asp Val Lys Ala Leu Val Lys Gly Cys Ala Ala Cys
            20                  25                  30
His Gly Val Lys Phe Glu Lys Lys Ala Leu Gly Lys Ser Lys Ile Val
        35                  40                  45
Asn Met Met Ser Glu Lys Glu Ile Glu Glu Asp Leu Met Ala Phe Lys
    50                  55                  60
Ser Gly Ala Asn Lys Asn Pro Val Met Thr Ala Gln Ala Lys Lys Leu
65                  70                  75                  80
Ser Asp Glu Asp Ile Lys Ala Leu Ala Lys Tyr Ile Pro Thr Leu Lys
                85                  90                  95
```

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 156 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...156

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

```
Met Arg Asp Phe Asn Asn Ile Gln Ile Thr Arg Leu Lys Val Arg Gln
1               5                  10                  15
Asn Ala Val Phe Glu Lys Leu Asp Leu Glu Phe Lys Asp Gly Leu Ser
            20                  25                  30
Ala Ile Ser Gly Ala Ser Gly Val Gly Lys Ser Val Leu Ile Ala Ser
        35                  40                  45
Leu Leu Gly Ala Phe Gly Leu Lys Glu Ser Asn Ala Ser Asn Ile Glu
```

```
    50                  55                  60
Val Glu Leu Ile Ala Pro Phe Leu Asp Thr Glu Glu Tyr Gly Ile Phe
65                  70                  75                  80
Arg Glu Asp Glu His Glu Pro Leu Val Ile Ser Val Ile Lys Lys Glu
                85                  90                  95
Lys Thr Arg Tyr Phe Leu Asn Gln Thr Ser Leu Ser Lys Asn Thr Leu
            100                 105                 110
Lys Ala Leu Leu Lys Gly Leu Ile Lys Arg Leu Ser Asn Asp Arg Phe
        115                 120                 125
Ser Gln Asn Glu Leu Asn Asp Ile Leu Met Leu Ser Leu Leu Asp Gly
    130                 135                 140
Tyr Ile Gln Asn Lys Asn Arg Arg Leu Ala Pro Phe
145                 150                 155
```

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 118 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...118

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

```
Val Met Leu Met Ala Ile Phe Thr Pro Tyr Ile Leu Ile Leu Lys Met
1               5                   10                  15
Met Lys Lys Ser Met Ser Leu Phe Ala Asn Met Gly Leu Glu Gln Ile
            20                  25                  30
Phe Cys Asn Arg Asp Ile Lys Asp Leu Asn Asp Phe Val Phe Gly Ile
        35                  40                  45
Glu Val Gly Leu Asp Ser Asn Ala Arg Lys Asn Arg Ser Arg Lys Ala
    50                  55                  60
Met Glu Asn His Leu Ile Gly Leu Phe Val Gln Ala Gln Leu Asn Phe
65                  70                  75                  80
Lys Glu Gln Val Asp Ile Arg Glu Phe Glu Asp Leu Arg Gln Ala Phe
                85                  90                  95
Gly Asn Asp Thr Lys Lys Phe Asp Phe Val Ile Phe Ser Lys Glu Lys
            100                 105                 110
Thr Tyr Phe His Arg Ser
        115
```

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 355 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...355

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

```
Met Asn Ile Lys Ile Leu Lys Ile Leu Val Gly Gly Leu Phe Phe Leu
1               5                  10                  15
Ser Leu Asn Ala His Leu Trp Gly Lys Gln Asp Asn Ser Phe Leu Gly
            20                  25                  30
Ile Gly Glu Arg Ala Tyr Lys Ser Gly Asn Tyr Ser Lys Ala Ala Ser
        35                  40                  45
Tyr Phe Lys Lys Ala Cys Asn Asp Gly Val Ser Glu Gly Cys Thr Gln
    50                  55                  60
Leu Gly Ile Ile Tyr Glu Asn Gly Gln Gly Thr Arg Ile Asp Tyr Lys
65                  70                  75                  80
Lys Ala Leu Glu Tyr Tyr Lys Thr Ala Cys Gln Ala Asp Asp Arg Glu
                85                  90                  95
Gly Cys Phe Gly Leu Gly Gly Leu Tyr Asp Glu Gly Leu Gly Thr Ala
            100                 105                 110
Gln Asn Tyr Gln Glu Ala Ile Asp Ala Tyr Ala Lys Ala Cys Val Leu
        115                 120                 125
Lys His Pro Glu Ser Cys Tyr Asn Leu Gly Ile Ile Tyr Asp Arg Lys
    130                 135                 140
Ile Lys Gly Asn Ala Ala Gln Ala Val Thr Tyr Tyr Gln Lys Ser Cys
145                 150                 155                 160
Asn Phe Asp Met Ala Lys Gly Cys Tyr Ile Leu Gly Thr Ala Tyr Glu
                165                 170                 175
Lys Gly Phe Leu Glu Val Lys Gln Ser Asn His Lys Ala Val Ile Tyr
            180                 185                 190
Tyr Leu Lys Ala Cys Arg Leu Asn Glu Gly Gln Ala Cys Arg Ala Leu
        195                 200                 205
Gly Ser Leu Phe Glu Asn Gly Asp Ala Gly Leu Asp Glu Asp Phe Glu
    210                 215                 220
Val Ala Phe Asp Tyr Leu Gln Lys Ala Cys Ala Leu Asn Asn Ser Gly
225                 230                 235                 240
Gly Cys Ala Ser Leu Gly Ser Met Tyr Met Leu Gly Arg Tyr Val Lys
                245                 250                 255
Lys Asp Pro Gln Lys Ala Phe Asn Tyr Phe Lys Gln Ala Cys Asp Met
            260                 265                 270
Gly Ser Ala Val Ser Cys Ser Arg Met Gly Phe Met Tyr Ser Gln Gly
        275                 280                 285
Asp Thr Val Ser Lys Asp Leu Arg Lys Ala Leu Asp Asn Tyr Glu Arg
    290                 295                 300
Gly Cys Asp Met Gly Asp Glu Val Gly Cys Phe Ala Leu Ala Gly Met
```

```
305                 310                 315                 320
Tyr Tyr Asn Met Lys Asp Lys Glu Asn Ala Ile Met Ile Tyr Asp Lys
                325                 330                 335
Gly Cys Lys Leu Gly Met Lys Gln Ala Cys Glu Asn Leu Thr Lys Leu
            340                 345                 350
Arg Gly Tyr
        355
```

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 193 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...193

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

```
Met Lys Glu Lys Asn Phe Trp Pro Leu Gly Ile Met Ser Val Leu Ile
1               5                   10                  15
Phe Gly Leu Gly Ile Val Val Phe Leu Val Val Phe Ala Leu Lys Asn
            20                  25                  30
Ser Pro Lys Asn Asp Leu Val Tyr Phe Lys Gly His Asn Glu Val Asp
        35                  40                  45
Leu Asn Phe Asn Ala Met Leu Lys Thr Tyr Glu Asn Phe Lys Ser Asn
    50                  55                  60
Tyr Arg Phe Ser Val Gly Leu Lys Pro Leu Thr Glu Ser Pro Lys Thr
65                  70                  75                  80
Pro Ile Leu Pro Tyr Phe Ser Lys Gly Thr His Gly Asp Lys Lys Ile
                85                  90                  95
Gln Glu Asn Leu Leu Asn Asn Ala Leu Ile Leu Glu Lys Ser Asn Thr
            100                 105                 110
Leu Tyr Ala Gln Leu Gln Pro Leu Lys Pro Ala Leu Asp Ser Pro Asn
        115                 120                 125
Ile Gln Val Tyr Leu Ala Phe Tyr Pro Ser Gln Ser Gln Pro Arg Leu
    130                 135                 140
Leu Gly Thr Leu Asp Cys Lys Asn Ala Cys Glu Pro Leu Lys Phe Asp
145                 150                 155                 160
Leu Leu Glu Gly Asp Lys Val Gly Arg Tyr Lys Ile Leu Phe Lys Phe
                165                 170                 175
Val Phe Lys Asn Lys Glu Glu Leu Ile Leu Glu Gln Leu Ala Phe Phe
            180                 185                 190
Lys
```

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 289 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...289

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

```
Leu Gly Ile Asn Met Cys Ser Lys Lys Ile Arg Asn Leu Ile Leu Cys
1               5                  10                  15
Phe Gly Phe Ile Leu Ser Leu Cys Ala Glu Glu Asn Ile Thr Lys Glu
            20                  25                  30
Asn Met Thr Glu Thr Asn Thr Thr Glu Glu Asn Thr Pro Lys Asp Ala
        35                  40                  45
Pro Ile Leu Leu Glu Glu Lys Arg Ala Gln Thr Leu Glu Leu Lys Glu
    50                  55                  60
Glu Asn Glu Val Ala Lys Lys Ile Asp Glu Lys Ser Leu Leu Glu Glu
65                  70                  75                  80
Ile His Lys Lys Lys Arg Gln Leu Tyr Met Leu Lys Gly Glu Leu His
                85                  90                  95
Glu Lys Asn Glu Ser Ile Leu Phe Gln Gln Met Ala Lys Asn Lys Ser
            100                 105                 110
Gly Phe Phe Ile Gly Val Ile Leu Gly Asp Ile Gly Ile Asn Ala Asn
        115                 120                 125
Pro Tyr Glu Lys Phe Glu Leu Leu Ser Asn Ile Gln Ala Ser Pro Leu
    130                 135                 140
Leu Tyr Gly Leu Arg Ser Gly Tyr Gln Lys Tyr Phe Ala Asn Gly Ile
145                 150                 155                 160
Ser Ala Leu Arg Phe Tyr Gly Glu Tyr Leu Gly Gly Ala Met Lys Gly
                165                 170                 175
Phe Lys Ser Asp Ser Leu Ala Ser Tyr Gln Thr Ala Ser Leu Asn Ile
            180                 185                 190
Asp Leu Leu Met Asp Lys Pro Ile Asp Lys Glu Lys Arg Phe Ala Leu
        195                 200                 205
Gly Ile Phe Gly Gly Val Gly Val Gly Trp Asn Gly Met Tyr Gln Asn
    210                 215                 220
Leu Lys Glu Ile Arg Gly Tyr Ser Gln Pro Asn Ala Phe Gly Leu Val
225                 230                 235                 240
Leu Asn Leu Gly Val Ser Met Thr Leu Asn Leu Lys His Arg Phe Glu
                245                 250                 255
Leu Ala Leu Lys Met Pro Pro Leu Lys Glu Thr Ser Gln Thr Phe Leu
            260                 265                 270
Tyr Tyr Phe Lys Ser Thr Asn Ile Tyr Tyr Ile Ser Tyr Asn Tyr Leu
```

275 280 285

Leu

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 668 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...668

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

Met Arg Lys Leu Phe Ile Pro Leu Leu Leu Phe Ser Ala Leu Glu Ala
1 5 10 15
Asn Glu Lys Asn Gly Phe Phe Ile Glu Ala Gly Phe Glu Thr Gly Leu
20 25 30
Leu Glu Gly Thr Gln Thr Gln Glu Lys Arg His Thr Thr Thr Lys Asn
35 40 45
Thr Tyr Ala Thr Tyr Asn Tyr Leu Pro Thr Asp Thr Ile Leu Lys Arg
50 55 60
Ala Ala Asn Leu Phe Thr Asn Ala Glu Ala Ile Ser Lys Leu Lys Phe
65 70 75 80
Ser Ser Leu Ser Pro Val Arg Val Leu Tyr Met Tyr Asn Gly Gln Leu
85 90 95
Thr Ile Glu Asn Phe Leu Pro Tyr Asn Leu Asn Asn Val Lys Leu Ser
100 105 110
Phe Thr Asp Ala Gln Gly Asn Thr Ile Asp Leu Gly Val Ile Glu Thr
115 120 125
Ile Pro Lys His Ser Lys Ile Val Leu Pro Gly Glu Ala Phe Asp Ser
130 135 140
Leu Lys Glu Ala Phe Asp Lys Ile Asp Pro Tyr Thr Leu Phe Leu Pro
145 150 155 160
Lys Phe Glu Ala Thr Ser Thr Ser Ile Ser Asp Thr Asn Thr Gln Arg
165 170 175
Val Phe Glu Thr Leu Asn Asn Ile Lys Thr Asn Leu Ile Met Lys Tyr
180 185 190
Ser Asn Glu Asn Pro Asn Asn Phe Asn Thr Cys Pro Tyr Asn Asn Asn
195 200 205
Gly Asn Thr Lys Asn Asp Cys Trp Gln Asn Phe Thr Pro Gln Thr Ala
210 215 220
Glu Glu Phe Thr Asn Leu Met Leu Asn Met Ile Ala Val Leu Asp Ser
225 230 235 240
Gln Ser Trp Gly Asp Ala Ile Leu Asn Ala Pro Phe Glu Phe Thr Asn

245 250 255
Ser Ser Thr Asp Cys Asp Ser Asp Pro Ser Lys Cys Val Asn Pro Gly
260 265 270
Val Asn Gly Arg Val Asp Thr Lys Val Asp Gln Gln Tyr Ile Leu Asn
275 280 285
Lys Gln Gly Ile Ile Asn Asn Phe Arg Lys Lys Ile Glu Ile Asp Ala
290 295 300
Val Val Leu Lys Asn Ser Gly Val Val Gly Leu Ala Asn Gly Tyr Gly
305 310 315 320
Asn Asp Gly Glu Tyr Gly Thr Leu Gly Val Glu Ala Tyr Ala Leu Asp
325 330 335
Pro Lys Lys Leu Phe Gly Asn Asp Leu Lys Thr Ile Asn Leu Glu Asp
340 345 350
Leu Arg Thr Ile Leu His Glu Phe Ser His Thr Lys Gly Tyr Gly His
355 360 365
Asn Gly Asn Met Thr Tyr Gln Arg Val Pro Val Thr Lys Asp Gly Gln
370 375 380
Val Glu Lys Asp Ser Asn Gly Lys Pro Lys Asp Ser Asp Gly Leu Pro
385 390 395 400
Tyr Asn Val Cys Ser Leu Tyr Gly Gly Ser Asn Gln Pro Ala Phe Pro
405 410 415
Ser Asn Tyr Pro Asn Ser Ile Tyr His Asn Cys Ala Asp Val Pro Ala
420 425 430
Gly Phe Leu Gly Val Thr Ala Ala Val Trp Gln Gln Leu Ile Asn Gln
435 440 445
Asn Ala Leu Pro Ile Asn Tyr Ala Asn Leu Gly Ser Gln Thr Asn Tyr
450 455 460
Asn Leu Asn Ala Ser Leu Asn Thr Gln Asp Leu Ala Asn Ser Met Leu
465 470 475 480
Ser Thr Ile Gln Lys Thr Phe Val Thr Ser Ser Val Thr Asn His His
485 490 495
Phe Ser Asn Ala Ser Gln Ser Phe Arg Ser Pro Ile Leu Gly Val Asn
500 505 510
Ala Lys Ile Gly Tyr Gln Asn Tyr Phe Asn Asp Phe Ile Gly Leu Ala
515 520 525
Tyr Tyr Gly Ile Ile Lys Tyr Asn Tyr Ala Lys Ala Val Asn Gln Lys
530 535 540
Val Gln Gln Leu Ser Tyr Gly Gly Gly Ile Asp Leu Leu Leu Asp Phe
545 550 555 560
Ile Thr Thr Tyr Ser Asn Lys Asn Ser Pro Thr Gly Ile Gln Thr Lys
565 570 575
Arg Asn Phe Ser Ser Ser Phe Gly Ile Phe Gly Gly Leu Arg Gly Leu
580 585 590
Tyr Asn Ser Tyr Tyr Val Leu Asn Lys Val Lys Gly Ser Gly Asn Leu
595 600 605
Asp Val Ala Thr Gly Leu Asn Tyr Arg Tyr Lys His Ser Lys Tyr Ser
610 615 620
Val Gly Ile Ser Ile Pro Leu Ile Gln Arg Lys Ala Ser Val Val Ser
625 630 635 640
Ser Gly Gly Asp Tyr Thr Asn Ser Phe Val Phe Asn Glu Gly Ala Ser
645 650 655
His Phe Lys Val Phe Phe Asn Tyr Gly Trp Val Phe
660 665

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...63

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

```
Met Asn Thr Glu Ile Leu Thr Ile Met Leu Val Val Ser Val Leu Met
1               5                  10                  15
Gly Leu Val Gly Leu Ile Ala Phe Leu Trp Gly Val Lys Ser Gly Gln
            20                  25                  30
Phe Asp Asp Glu Lys Arg Met Leu Glu Ser Val Leu Tyr Asp Ser Ala
        35                  40                  45
Ser Asp Leu Asn Glu Ala Ile Leu Gln Glu Lys Arg Gln Lys Asn
    50                  55                  60
```

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 406 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...406

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

```
Met Val Phe Phe His Lys Lys Ile Ile Leu Asn Phe Ile Tyr Ser Leu
1               5                  10                  15
Met Val Ala Phe Leu Phe His Leu Ser Tyr Gly Val Leu Leu Lys Ala
            20                  25                  30
Asp Gly Met Ala Lys Lys Gln Thr Leu Leu Val Gly Glu Arg Leu Val
        35                  40                  45
Trp Asp Lys Leu Thr Leu Leu Gly Phe Leu Glu Lys Asn His Ile Pro
```

```
     50                  55                  60
Gln Lys Leu Tyr Tyr Asn Leu Ser Ser Gln Asp Lys Glu Leu Ser Ala
65                  70                  75                  80
Glu Ile Gln Ser Asn Val Thr Tyr Tyr Thr Leu Arg Asp Ala Asn Asn
                85                  90                  95
Thr Leu Ile Gln Ala Leu Ile Pro Ile Ser Gln Asp Leu Gln Ile His
            100                 105                 110
Ile Tyr Lys Lys Gly Glu Asp Tyr Phe Leu Asp Phe Ile Pro Ile Val
        115                 120                 125
Phe Thr Arg Lys Glu Arg Thr Leu Leu Leu Ser Leu Gln Thr Ser Pro
    130                 135                 140
Tyr Gln Asp Ile Val Lys Ala Thr Asn Asp Pro Leu Leu Ala Asn Gln
145                 150                 155                 160
Leu Met Asn Ala Tyr Lys Lys Ser Val Pro Phe Lys Arg Leu Val Lys
                165                 170                 175
Asn Asp Lys Ile Ala Ile Val Tyr Thr Arg Asp Tyr Arg Val Gly Gln
            180                 185                 190
Ala Phe Gly Gln Pro Thr Ile Lys Met Ala Met Val Ser Ser Arg Leu
        195                 200                 205
His Gln Tyr Tyr Leu Phe Ser His Ser Asn Gly Arg Tyr Tyr Asp Ser
    210                 215                 220
Lys Ala Gln Glu Val Ala Gly Phe Leu Leu Glu Thr Pro Val Lys Tyr
225                 230                 235                 240
Thr Arg Ile Ser Ser Pro Phe Ser Tyr Gly Arg Phe His Pro Val Leu
                245                 250                 255
Lys Val Lys Arg Pro His Tyr Gly Val Asp Tyr Ala Ala Lys His Gly
            260                 265                 270
Ser Leu Ile His Ser Ala Ser Asp Gly Arg Val Gly Phe Ile Gly Val
        275                 280                 285
Lys Ala Gly Tyr Gly Lys Val Val Glu Ile His Leu Asn Glu Leu Arg
    290                 295                 300
Leu Val Tyr Ala His Met Ser Ala Phe Ala Asn Gly Leu Lys Lys Gly
305                 310                 315                 320
Ser Phe Val Lys Lys Gly Gln Ile Ile Gly Arg Val Gly Ser Thr Gly
                325                 330                 335
Leu Ser Thr Gly Pro His Leu His Phe Gly Val Tyr Lys Asn Ser Arg
            340                 345                 350
Pro Ile Asn Pro Leu Gly Tyr Ile Arg Thr Ala Lys Ser Lys Leu His
        355                 360                 365
Gly Lys Gln Arg Glu Val Phe Leu Glu Lys Ala Gln Tyr Ser Lys Gln
    370                 375                 380
Lys Leu Glu Glu Leu Phe Lys Thr His Ser Phe Glu Lys Asn Ser Phe
385                 390                 395                 400
Tyr Leu Leu Glu Gly Phe
                405
```

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 296 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...296

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

```
Leu Phe Leu Val Lys Lys Ile Gly Val Val Ile Met Ile Leu Val Cys
1               5                  10                  15
Phe Leu Ala Cys Ser Gln Glu Ser Phe Ile Lys Met Gln Lys Lys Ala
            20                  25                  30
Gln Glu Gln Glu Asn Asp Gly Ser Lys Arg Pro Ser Tyr Val Asp Ser
        35                  40                  45
Asp Tyr Glu Val Phe Ser Glu Thr Ile Phe Leu Gln Asn Met Val Tyr
    50                  55                  60
Gln Pro Ile Glu Glu Arg Asn Ala Phe Phe Gln Leu Thr Lys Asp Glu
65                  70                  75                  80
Asp Asn Ser Phe Asn Pro Glu Asn Ser Val Ile Leu Leu Asn Glu Pro
                85                  90                  95
Ser Asp Asn Ser Glu Lys Asn Leu Leu Ser Tyr Pro Asn Asp Pro Asn
            100                 105                 110
Asn Asn Glu Asp Asn Ala Asn Asn Ser Gln Lys Asn Pro Phe Leu Tyr
        115                 120                 125
Lys Pro Lys Arg Lys Thr Lys Asn Pro Lys Leu Ile Glu Tyr Ser Gln
    130                 135                 140
Gln Asp Phe Tyr Pro Leu Lys Asn Gly Asp Ile Ile Met Ser Lys Glu
145                 150                 155                 160
Gly Asp Gln Trp Leu Ile Glu Ile Gln Ser Lys Ala Leu Lys Arg Phe
                165                 170                 175
Leu Lys Asp Gln Asn Asp Lys Asp Arg Gln Ile Gln Thr Phe Thr Phe
            180                 185                 190
Asn Asp Thr Lys Thr Gln Ile Ala Gln Ile Lys Gly Lys Ile Ser Ser
        195                 200                 205
Tyr Val Tyr Thr Thr Asn Asn Gly Ser Leu Ser Leu Arg Pro Phe Tyr
    210                 215                 220
Glu Ser Phe Leu Leu Glu Lys Lys Ser Asp Asn Val Tyr Thr Ile Glu
225                 230                 235                 240
Asn Lys Ala Leu Asp Thr Met Glu Ile Ser Lys Cys Gln Met Val Leu
                245                 250                 255
Lys Lys His Ser Thr Asp Lys Leu Asp Ser Gln His Lys Ala Ile Ser
            260                 265                 270
Ile Asp Leu Asp Phe Lys Lys Glu Arg Phe Lys Ser Asp Thr Glu Leu
        275                 280                 285
Phe Leu Glu Cys Leu Lys Glu Ser
    290                 295
```

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 248 amino acids
(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...248

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

```
Val Ser Tyr Asp Asn Thr Asp Asp Tyr Tyr Phe Pro Arg Asn Gly Val
1               5                  10                  15
Ile Phe Ser Ser Tyr Ala Thr Met Ser Gly Leu Pro Ser Ser Gly Thr
            20                  25                  30
Leu Asn Ser Trp Asn Gly Leu Gly Gly Asn Val Arg Asn Thr Lys Val
        35                  40                  45
Tyr Gly Lys Phe Ala Ala Tyr His His Leu Gln Lys Tyr Leu Leu Ile
    50                  55                  60
Asp Leu Ile Ala Arg Phe Lys Thr Gln Gly Gly Tyr Ile Phe Arg Tyr
65                  70                  75                  80
Asn Thr Asp Asp Tyr Leu Pro Leu Asn Ser Thr Phe Tyr Met Gly Gly
                85                  90                  95
Val Thr Thr Val Arg Gly Phe Arg Asn Gly Ser Ile Thr Pro Lys Asp
            100                 105                 110
Glu Phe Gly Leu Trp Leu Gly Gly Asp Gly Ile Phe Thr Ala Ser Thr
        115                 120                 125
Glu Leu Ser Tyr Gly Val Leu Lys Ala Ala Lys Met Arg Leu Ala Trp
    130                 135                 140
Phe Phe Asp Phe Gly Phe Leu Thr Phe Lys Thr Pro Thr Arg Gly Ser
145                 150                 155                 160
Phe Phe Tyr Asn Ala Pro Thr Thr Thr Ala Asn Phe Lys Asp Tyr Gly
                165                 170                 175
Val Val Gly Ala Gly Phe Glu Arg Ala Thr Trp Arg Ala Ser Thr Gly
            180                 185                 190
Leu Gln Ile Glu Trp Ile Ser Pro Met Gly Pro Leu Val Leu Ile Phe
        195                 200                 205
Pro Ile Ala Phe Phe Asn Gln Trp Gly Asp Gly Asn Gly Lys Lys Cys
    210                 215                 220
Lys Gly Leu Cys Phe Asn Pro Asn Met Asn Asp Tyr Thr Gln His Phe
225                 230                 235                 240
Glu Phe Ser Met Gly Thr Arg Phe
                245
```

(2) INFORMATION FOR SEQ ID NO:113:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 335 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...335

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

```
Val Gln His Phe Asn Phe Leu Tyr Lys Asp Ser Leu Phe Ser Ile Ala
1               5                  10                  15
Leu Phe Thr Phe Ile Ile Ala Leu Val Ile Leu Leu Glu Gln Ala Arg
            20                  25                  30
Ala Tyr Phe Thr Arg Lys Arg Asn Lys Lys Phe Leu Gln Lys Phe Ala
        35                  40                  45
Gln Asn Gln Asn Ala Tyr Ala Ser Ser Glu Asn Leu Asp Glu Leu Leu
    50                  55                  60
Lys His Ala Lys Ile Ser Ser Leu Met Phe Leu Ala Arg Ala Tyr Ser
65                  70                  75                  80
Lys Ala Asp Val Glu Met Ser Ile Glu Ile Leu Lys Gly Leu Leu Asn
                85                  90                  95
Arg Pro Leu Lys Asp Glu Glu Lys Ile Ala Val Leu Asp Leu Leu Ala
            100                 105                 110
Lys Asn Tyr Phe Ser Val Gly Tyr Leu Gln Lys Thr Lys Asp Thr Val
        115                 120                 125
Lys Glu Ile Leu Arg Phe Ser Pro Arg Asn Val Glu Ala Leu Leu Lys
    130                 135                 140
Leu Leu His Ala Tyr Glu Leu Glu Lys Asp Tyr Ser Lys Ala Leu Glu
145                 150                 155                 160
Thr Leu Glu Cys Leu Glu Glu Leu Glu Val Pro Lys Ile Glu Thr Ile
                165                 170                 175
Lys Asn Tyr Leu Tyr Leu Met His Leu Ile Glu Asn Lys Glu Asp Ala
            180                 185                 190
Ala Lys Ile Leu His Val Ser Lys Ala Ser Leu Asp Leu Lys Lys Ile
        195                 200                 205
Ala Leu Asn His Leu Lys Ser His Asp Glu Asn Leu Phe Trp Gln Glu
    210                 215                 220
Ile Asp Thr Thr Glu Arg Leu Glu Asn Val Ile Asp Leu Leu Trp Asp
225                 230                 235                 240
Met Asn Ile Pro Ala Phe Ile Leu Glu Lys His Ala Leu Leu Gln Asp
                245                 250                 255
Ile Ala Arg Ser Gln Gly Leu Leu Leu Asp His Lys Pro Cys Gln Ile
            260                 265                 270
Phe Glu Leu Glu Val Leu Arg Ala Leu Leu His Ser Pro Ile Lys Ala
        275                 280                 285
Ser Leu Thr Phe Glu Tyr Arg Cys Lys His Cys Lys Gln Ile Phe Pro
    290                 295                 300
Phe Glu Ser His Arg Cys Pro Val Cys Tyr Gln Leu Ala Phe Met Asp
305                 310                 315                 320
Met Val Leu Lys Ile Ser Lys Lys Thr His Ala Met Gly Val Asp
                325                 330                 335
```

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 413 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...413

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

```
Met Arg Lys Ile Phe Ser Tyr Ile Ser Lys Val Leu Leu Phe Ile Gly
1               5                  10                  15
Val Val Tyr Ala Glu Pro Asp Ser Lys Val Glu Ala Leu Glu Gly Arg
            20                  25                  30
Lys Gln Glu Ser Ser Leu Asp Lys Lys Ile Arg Gln Glu Leu Lys Ser
        35                  40                  45
Lys Glu Leu Lys Asn Lys Glu Leu Lys Asn Lys Asp Leu Lys Asn Lys
    50                  55                  60
Glu Glu Lys Lys Glu Thr Lys Ala Lys Arg Lys Pro Arg Ala Glu Val
65                  70                  75                  80
His His Gly Asp Ala Lys Asn Pro Thr Pro Lys Ile Thr Pro Pro Lys
                85                  90                  95
Ile Lys Gly Ser Ser Lys Gly Val Gln Asn Gln Gly Val Gln Asn Asn
            100                 105                 110
Ala Pro Lys Pro Glu Glu Lys Asp Thr Thr Pro Gln Ala Thr Glu Lys
        115                 120                 125
Asn Lys Glu Thr Ser Pro Ser Ser Gln Phe Asn Ser Ile Phe Gly Asn
    130                 135                 140
Pro Asn Asn Ala Thr Asn Asn Thr Leu Glu Asp Lys Val Val Gly Gly
145                 150                 155                 160
Ile Ser Leu Leu Val Asn Gly Ser Pro Ile Thr Leu Tyr Gln Ile Gln
                165                 170                 175
Glu Glu Gln Glu Lys Ser Lys Val Ser Lys Ala Gln Ala Arg Asp Arg
            180                 185                 190
Leu Ile Ala Glu Arg Ile Lys Asn Gln Glu Ile Glu Arg Leu Lys Ile
        195                 200                 205
His Val Asp Asp Asp Lys Leu Asp Gln Glu Met Ala Met Met Ala Gln
    210                 215                 220
Gln Gln Gly Met Asp Leu Asp His Phe Lys Gln Met Leu Met Ala Glu
225                 230                 235                 240
Gly His Tyr Lys Leu Tyr Arg Asp Gln Leu Lys Glu His Leu Glu Met
                245                 250                 255
Gln Glu Leu Leu Arg Asn Ile Leu Leu Thr Asn Val Asp Thr Ser Ser
            260                 265                 270
```

```
Glu Thr Lys Met Arg Glu Tyr Tyr Asn Lys His Lys Glu Gln Phe Ser
        275                 280                 285
Ile Pro Thr Glu Ile Glu Thr Val Arg Tyr Thr Ser Thr Asn Gln Glu
    290                 295                 300
Asp Leu Glu Arg Ala Met Ala Asp Pro Asn Leu Glu Val Pro Gly Val
305                 310                 315                 320
Ser Lys Ala Asn Glu Lys Ile Glu Met Lys Thr Leu Asn Pro Gln Ile
                325                 330                 335
Ala Gln Val Phe Ile Ser His Glu Gln Gly Ser Phe Thr Pro Val Met
            340                 345                 350
Asn Gly Gly Gly Gly Gln Phe Ile Thr Phe Tyr Ile Lys Glu Lys Arg
        355                 360                 365
Gly Lys Asn Glu Val Ser Phe Ser Gln Ala Lys Gln Phe Ile Ala Gln
    370                 375                 380
Lys Leu Val Glu Glu Ser Lys Asp Lys Ile Leu Glu Glu His Phe Glu
385                 390                 395                 400
Lys Leu Arg Val Lys Ser Arg Ile Val Met Ile Arg Glu
                405                 410
```

(2) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 186 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...186

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

```
Met Ile Lys Arg Ile Ala Cys Ile Leu Ser Leu Ser Ala Ser Leu Ala
1               5                   10                  15
Leu Ala Gly Glu Val Asn Gly Phe Phe Met Gly Ala Gly Tyr Gln Gln
            20                  25                  30
Gly Arg Tyr Gly Pro Tyr Asn Ser Asn Tyr Ser Asp Trp Arg His Gly
        35                  40                  45
Asn Asp Leu Tyr Gly Leu Asn Phe Lys Leu Gly Phe Val Gly Phe Ala
    50                  55                  60
Asn Lys Trp Phe Gly Ala Arg Val Tyr Gly Phe Leu Asp Trp Phe Asn
65                  70                  75                  80
Thr Ser Gly Thr Glu His Thr Lys Thr Asn Leu Leu Thr Tyr Gly Gly
                85                  90                  95
Gly Gly Asp Leu Ile Val Asn Leu Ile Pro Leu Asp Lys Phe Ala Leu
            100                 105                 110
Gly Leu Ile Gly Gly Val Gln Leu Ala Gly Asn Thr Trp Met Phe Pro
        115                 120                 125
```

Tyr Asp Val Asn Gln Thr Arg Phe Gln Phe Leu Trp Asn Leu Gly Gly
130 135 140
Arg Met Arg Val Gly Asp Arg Ser Ala Phe Glu Ala Gly Val Lys Phe
145 150 155 160
Pro Met Val Asn Gln Gly Ser Lys Asp Val Gly Leu Ile Arg Tyr Tyr
165 170 175
Ser Trp Tyr Val Asp Tyr Val Phe Thr Phe
180 185

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 242 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...242

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

Met Lys Lys Phe Phe Ser Gln Ser Leu Leu Ala Leu Ile Ile Ser Met
1 5 10 15
Asn Ala Val Ser Gly Met Asp Gly Asn Gly Val Phe Leu Gly Ala Gly
20 25 30
Tyr Leu Gln Gly Gln Ala Gln Met His Ala Asp Ile Asn Ser Gln Lys
35 40 45
Gln Ala Thr Asn Ala Thr Ile Lys Gly Phe Asp Ala Leu Leu Gly Tyr
50 55 60
Gln Phe Phe Phe Glu Lys His Phe Gly Leu Arg Leu Tyr Gly Phe Phe
65 70 75 80
Asp Tyr Ala His Ala Asn Ser Ile Lys Leu Lys Asn Pro Asn Tyr Asn
85 90 95
Ser Glu Ala Ala Gln Val Ala Ser Gln Ile Leu Gly Lys Gln Glu Ile
100 105 110
Asn Arg Leu Thr Asn Ile Ala Asp Pro Arg Thr Phe Glu Pro Asn Met
115 120 125
Leu Thr Tyr Gly Gly Ala Met Asp Val Met Val Asn Val Ile Asn Asn
130 135 140
Gly Ile Met Ser Leu Gly Ala Phe Gly Gly Ile Gln Leu Ala Gly Asn
145 150 155 160
Ser Trp Leu Met Ala Thr Pro Ser Phe Glu Gly Ile Leu Val Glu Gln
165 170 175
Ala Leu Val Ser Lys Lys Ala Thr Ser Phe Gln Phe Leu Phe Asn Val
180 185 190
Gly Ala Arg Leu Arg Ile Leu Lys His Ser Ser Ile Glu Ala Gly Val
195 200 205

```
Lys Phe Pro Met Leu Lys Lys Asn Pro Tyr Ile Thr Ala Lys Asn Leu
    210                 215                 220
Asp Ile Gly Phe Arg Arg Val Tyr Ser Trp Tyr Val Asn Tyr Val Phe
225                 230                 235                 240
Thr Phe
```

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 256 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...256

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

```
Met Gly Tyr Ala Ser Lys Leu Ala Leu Lys Ile Cys Leu Val Gly Leu
1               5                  10                  15
Cys Leu Phe Ser Thr Leu Gly Ala Glu His Leu Glu Gln Lys Gly Asn
            20                  25                  30
Tyr Ile Tyr Lys Gly Glu Glu Ala Tyr Asn Asn Lys Glu Tyr Glu Arg
        35                  40                  45
Ala Ala Ser Phe Tyr Lys Ser Ala Ile Lys Asn Gly Glu Ser Leu Ala
    50                  55                  60
Tyr Ile Leu Leu Gly Ile Met Tyr Glu Asn Gly Arg Gly Val Pro Lys
65                  70                  75                  80
Asp Tyr Lys Lys Ala Val Glu Tyr Phe Gln Lys Ala Val Asp Asn Asp
                85                  90                  95
Ile Pro Arg Gly Tyr Asn Asn Leu Gly Val Met Tyr Lys Glu Gly Lys
            100                 105                 110
Gly Val Pro Lys Asp Glu Lys Lys Ala Val Glu Tyr Phe Arg Ile Ala
        115                 120                 125
Thr Glu Lys Gly Tyr Thr Asn Ala Tyr Ile Asn Leu Gly Ile Met Tyr
    130                 135                 140
Met Glu Gly Arg Gly Val Pro Ser Asn Tyr Ala Lys Ala Thr Glu Cys
145                 150                 155                 160
Phe Arg Lys Ala Met His Lys Gly Asn Val Glu Ala Tyr Ile Leu Leu
                165                 170                 175
Gly Asp Ile Tyr Tyr Ser Gly Asn Asp Gln Leu Gly Ile Glu Pro Asp
            180                 185                 190
Lys Asp Lys Ala Val Val Tyr Tyr Lys Met Ala Ala Asp Val Ser Ser
        195                 200                 205
Ser Arg Ala Tyr Glu Gly Leu Ser Glu Ser Tyr Arg Tyr Gly Leu Gly
    210                 215                 220
```

```
Val Glu Lys Asp Lys Lys Lys Ala Glu Glu Tyr Met Gln Lys Ala Cys
225                 230                 235                 240
Asp Phe Asp Ile Asp Lys Asn Cys Lys Lys Lys Asn Thr Ser Ser Arg
                245                 250                 255
```

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 657 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...657

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

```
Met Arg Lys Leu Phe Ile Pro Leu Leu Leu Phe Ser Ala Leu Glu Ala
1               5                  10                  15
Asn Glu Lys Asn Gly Phe Phe Ile Glu Ala Gly Phe Glu Thr Gly Leu
            20                  25                  30
Leu Glu Gly Thr Gln Thr Gln Glu Lys Arg His Thr Thr Thr Lys Asn
        35                  40                  45
Thr Tyr Ala Thr Tyr Asn Tyr Leu Pro Thr Asp Thr Ile Leu Lys Arg
    50                  55                  60
Ala Ala Asn Leu Phe Thr Asn Ala Glu Ala Ile Ser Lys Leu Lys Phe
65                  70                  75                  80
Ser Ser Leu Ser Pro Val Arg Val Leu Tyr Met Tyr Asn Gly Gln Leu
                85                  90                  95
Thr Ile Glu Asn Phe Leu Pro Tyr Asn Leu Asn Asn Val Lys Leu Ser
            100                 105                 110
Phe Thr Asp Ala Gln Gly Asn Val Ile Asp Leu Gly Val Ile Glu Thr
        115                 120                 125
Ile Pro Lys His Ser Lys Ile Val Leu Pro Gly Glu Ala Phe Asp Ser
    130                 135                 140
Leu Lys Ile Asp Pro Tyr Thr Leu Phe Leu Pro Lys Ile Glu Ala Thr
145                 150                 155                 160
Ser Thr Ser Ile Ser Asp Ala Asn Thr Gln Arg Val Phe Glu Thr Leu
                165                 170                 175
Asn Lys Ile Lys Thr Asn Leu Val Val Asn Tyr Arg Asn Glu Asn Lys
            180                 185                 190
Phe Lys Asp His Glu Asn His Trp Glu Ala Phe Thr Pro Gln Thr Ala
        195                 200                 205
Glu Glu Phe Thr Asn Leu Met Leu Asn Met Ile Ala Val Leu Asp Ser
    210                 215                 220
Gln Ser Trp Gly Asp Ala Ile Leu Asn Ala Pro Phe Glu Phe Thr Asn
225                 230                 235                 240
```

Ser Pro Thr Asp Cys Asp Asn Asp Pro Ser Lys Cys Val Asn Pro Gly
245 250 255
Thr Asn Gly Leu Val Asn Ser Lys Val Asp Gln Lys Tyr Val Leu Asn
260 265 270
Lys Gln Asp Ile Val Asn Lys Phe Lys Asn Lys Ala Asp Leu Asp Val
275 280 285
Ile Val Leu Lys Asp Ser Gly Val Val Gly Leu Gly Ser Asp Ile Thr
290 295 300
Pro Ser Asn Asn Asp Asp Gly Lys His Tyr Gly Gln Leu Gly Val Val
305 310 315 320
Ala Ser Ala Leu Asp Pro Lys Lys Leu Phe Gly Asp Asn Leu Lys Thr
325 330 335
Ile Asn Leu Glu Asp Leu Arg Thr Ile Leu His Glu Phe Ser His Thr
340 345 350
Lys Gly Tyr Gly His Asn Gly Asn Met Thr Tyr Gln Arg Val Pro Val
355 360 365
Thr Lys Asp Gly Gln Val Glu Lys Asp Ser Asn Gly Lys Pro Lys Asp
370 375 380
Ser Asp Gly Leu Pro Tyr Asn Val Cys Ser Leu Tyr Gly Gly Ser Asn
385 390 395 400
Gln Pro Ala Phe Pro Ser Asn Tyr Pro Asn Ser Ile Tyr His Asn Cys
405 410 415
Ala Asp Val Pro Ala Gly Phe Leu Gly Val Thr Ala Ala Val Trp Gln
420 425 430
Gln Leu Ile Asn Gln Asn Ala Leu Pro Ile Asn Tyr Ala Asn Leu Gly
435 440 445
Ser Gln Thr Asn Tyr Asn Leu Asn Ala Ser Leu Asn Thr Gln Asp Leu
450 455 460
Ala Asn Ser Met Leu Ser Thr Ile Gln Lys Thr Phe Val Thr Ser Ser
465 470 475 480
Val Thr Asn His His Phe Ser Asn Ala Ser Gln Ser Phe Arg Ser Pro
485 490 495
Ile Leu Gly Val Asn Ala Lys Ile Gly Tyr Gln Asn Tyr Phe Asn Asp
500 505 510
Phe Ile Gly Leu Ala Tyr Tyr Gly Ile Ile Lys Tyr Asn Tyr Ala Lys
515 520 525
Ala Val Asn Gln Lys Val Gln Gln Leu Ser Tyr Gly Gly Gly Ile Asp
530 535 540
Leu Leu Leu Asp Phe Ile Thr Thr Tyr Ser Asn Lys Asn Ser Pro Thr
545 550 555 560
Gly Ile Gln Thr Lys Arg Asn Phe Ser Ser Ser Phe Gly Ile Phe Gly
565 570 575
Gly Leu Arg Gly Leu Tyr Asn Ser Tyr Tyr Val Leu Asn Lys Val Lys
580 585 590
Gly Ser Gly Asn Leu Asp Val Ala Thr Gly Leu Asn Tyr Arg Tyr Lys
595 600 605
His Ser Lys Tyr Ser Val Gly Ile Ser Ile Pro Leu Ile Gln Arg Lys
610 615 620
Ala Ser Val Val Ser Ser Gly Gly Asp Tyr Thr Asn Ser Phe Val Phe
625 630 635 640
Asn Glu Gly Ala Ser His Phe Lys Val Phe Phe Asn Tyr Gly Trp Val
645 650 655
Phe

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 167 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...167

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

```
Met Lys Leu Val Ser Leu Ile Val Ala Leu Val Phe Cys Cys Phe Leu
1               5                  10                  15
Gly Ala Val Glu Leu Pro Gly Val Tyr Gln Thr Gln Glu Phe Leu Tyr
            20                  25                  30
Met Lys Ser Ser Phe Val Glu Phe Phe Glu His Asn Gly Lys Phe Tyr
        35                  40                  45
Ala Tyr Gly Ile Ser Asp Val Asp Gly Ser Lys Ala Lys Lys Asp Lys
    50                  55                  60
Leu Asn Pro Asn Pro Lys Leu Arg Asn Arg Ser Asp Lys Gly Val Val
65                  70                  75                  80
Phe Leu Ser Asp Leu Ile Lys Val Gly Glu Gln Ser Tyr Lys Gly Gly
                85                  90                  95
Lys Ala Tyr Asn Phe Tyr Asp Gly Lys Thr Tyr His Val Arg Val Thr
            100                 105                 110
Gln Asn Ser Asn Gly Asp Leu Glu Phe Thr Ser Ser Tyr Asp Lys Trp
        115                 120                 125
Gly Tyr Val Gly Lys Thr Phe Thr Trp Lys Arg Leu Ser Asp Glu Glu
    130                 135                 140
Ile Lys Asn Leu Lys Leu Lys Arg Phe Asn Leu Asp Glu Val Leu Lys
145                 150                 155                 160
Thr Leu Lys Asp Ser Pro Ile
                165
```

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 294 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...294

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

```
Met Ser Asn Gln Ala Ser His Leu Asp Asn Phe Met Asn Ala Lys Asn
1               5                  10                  15
Pro Lys Ser Phe Phe Asp Asn Lys Gly Asn Thr Lys Phe Ile Ala Ile
            20                  25                  30
Thr Ser Gly Lys Gly Gly Val Gly Lys Ser Asn Ile Ser Ala Asn Leu
        35                  40                  45
Ala Tyr Ser Leu Tyr Lys Lys Gly Tyr Lys Val Gly Val Phe Asp Ala
    50                  55                  60
Asp Ile Gly Leu Ala Asn Leu Asp Val Ile Phe Gly Val Lys Thr His
65                  70                  75                  80
Lys Asn Ile Leu His Ala Leu Lys Gly Glu Ala Lys Leu Gln Glu Ile
                85                  90                  95
Ile Cys Glu Ile Glu Pro Gly Leu Cys Leu Ile Pro Gly Asp Ser Gly
            100                 105                 110
Glu Glu Ile Leu Lys Tyr Ile Ser Gly Ala Glu Ala Leu Asp Arg Phe
        115                 120                 125
Val Asp Glu Glu Gly Val Leu Ser Ser Leu Asp Tyr Ile Val Ile Asp
    130                 135                 140
Thr Gly Ala Gly Ile Gly Ala Thr Thr Gln Ala Phe Leu Asn Ala Ser
145                 150                 155                 160
Asp Cys Val Val Ile Val Thr Thr Pro Asp Pro Ser Ala Ile Thr Asp
                165                 170                 175
Ala Tyr Ala Cys Ile Lys Ile Asn Ser Lys Asn Lys Asp Glu Leu Phe
            180                 185                 190
Leu Ile Ala Asn Met Val Ala Gln Pro Lys Glu Gly Arg Ala Thr Tyr
        195                 200                 205
Glu Arg Leu Phe Lys Val Ala Lys Asn Asn Ile Ala Ser Leu Glu Leu
    210                 215                 220
His Tyr Leu Gly Ala Ile Glu Asn Ser Ser Leu Leu Lys Arg Tyr Val
225                 230                 235                 240
Arg Glu Arg Lys Ile Leu Arg Lys Ile Ala Pro Asn Asp Leu Phe Ser
                245                 250                 255
Gln Ser Ile Asp Gln Ile Ala Ser Leu Leu Val Ser Lys Leu Glu Thr
            260                 265                 270
Gly Thr Leu Glu Ile Pro Lys Glu Gly Leu Lys Ser Phe Phe Lys Arg
        275                 280                 285
Leu Leu Lys Tyr Leu Gly
    290
```

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...372

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

```
Leu Glu Pro Ser Arg Asn Arg Leu Lys His Ala Ala Phe Phe Val Gly
1               5                  10                  15
Leu Phe Ile Val Leu Phe Leu Ile Ile Met Lys His Gln Thr Ser Pro
            20                  25                  30
Tyr Ala Phe Thr His Asn Gln Ala Leu Val Thr Gln Thr Pro Pro Tyr
        35                  40                  45
Phe Thr Gln Leu Thr Ile Pro Lys Pro Asn Asp Ala Leu Ser Ala His
    50                  55                  60
Ala Ser Ser Leu Ile Ser Leu Pro Asn Asp Asn Leu Leu Ser Ala Tyr
65                  70                  75                  80
Phe Ser Gly Thr Lys Glu Gly Ala Arg Asp Val Lys Ile Ser Ala Asn
                85                  90                  95
Leu Phe Asp Ser Lys Thr Asn Arg Trp Ser Glu Ala Phe Ile Leu Leu
            100                 105                 110
Thr Lys Glu Glu Leu Ser His His Ser His Glu Tyr Ile Lys Lys Leu
        115                 120                 125
Gly Asn Pro Leu Leu Phe Leu His Asp Asn Lys Ile Leu Leu Phe Val
    130                 135                 140
Val Gly Val Ser Met Gly Gly Trp Ala Thr Ser Lys Ile Tyr Gln Phe
145                 150                 155                 160
Glu Ser Ala Leu Glu Pro Ile His Phe Lys Phe Ala Arg Lys Leu Ser
                165                 170                 175
Leu Ser Pro Phe Leu Asn Leu Ser His Leu Val Arg Asn Lys Pro Leu
            180                 185                 190
Asn Thr Thr Asp Gly Gly Phe Met Leu Pro Leu Tyr His Glu Leu Ala
        195                 200                 205
Thr Gln Tyr Pro Leu Leu Leu Lys Phe Asp Gln Gln Asn Asn Pro Arg
    210                 215                 220
Glu Leu Leu Arg Pro Asn Thr Leu Asn His Gln Leu Gln Pro Ser Leu
225                 230                 235                 240
Thr Pro Phe Lys Asp Cys Ala Val Met Ala Phe Arg Asn His Ser Phe
                245                 250                 255
Lys Asp Ser Leu Met Leu Glu Thr Cys Lys Thr Pro Thr Asp Trp Gln
            260                 265                 270
Lys Pro Ile Ser Thr Asn Leu Lys Asn Leu Asp Asp Ser Leu Asn Leu
        275                 280                 285
Leu Asn Leu Asn Gly Ile Leu Tyr Leu Ile His Asn Pro Ser Asp Leu
    290                 295                 300
Ser Leu Arg Arg Lys Glu Leu Trp Leu Ser Lys Leu Glu Asn Ser Asn
305                 310                 315                 320
Ser Phe Lys Thr Leu Lys Val Leu Asp Lys Ala Asn Glu Val Ser Tyr
                325                 330                 335
Pro Ser Tyr Ser Leu Asn Pro His Phe Ile Asp Ile Val Tyr Thr Tyr
```

```
            340                 345                 350
Asn Arg Ser His Ile Lys His Ile Arg Phe Asn Met Ala Tyr Leu Asn
        355                 360                 365
Ser Leu Leu Lys
    370
```

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 978 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...978

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

```
Met Lys Lys Arg Lys His Val Ser Lys Lys Val Phe Asn Val Ile Ile
1               5                  10                  15
Leu Phe Val Ala Val Phe Thr Leu Leu Val Val Ile His Lys Thr Leu
            20                  25                  30
Ser Asn Gly Ile His Ile Gln Asn Leu Lys Ile Gly Lys Leu Gly Ile
        35                  40                  45
Ser Glu Leu Tyr Leu Lys Leu Asn Asn Lys Leu Ser Leu Glu Val Glu
    50                  55                  60
Arg Val Asp Leu Ser Ser Phe Phe His Gln Lys Pro Thr Lys Lys Arg
65                  70                  75                  80
Leu Glu Val Ser Asp Leu Ile Lys Asn Ile Arg Tyr Gly Ile Trp Ala
                85                  90                  95
Val Ser Tyr Phe Glu Lys Leu Lys Val Lys Glu Ile Ile Leu Asp Asp
            100                 105                 110
Lys Asn Lys Ala Asn Ile Phe Phe Asp Gly Asn Lys Tyr Glu Leu Glu
        115                 120                 125
Phe Pro Gly Ile Lys Gly Glu Phe Ser Leu Glu Asp Asp Lys Asn Ile
    130                 135                 140
Lys Leu Lys Ile Ile Asn Leu Leu Phe Lys Asp Val Lys Val Gln Val
145                 150                 155                 160
Asp Gly Asn Ala His Tyr Ser Pro Lys Ala Arg Lys Met Ala Phe Asn
                165                 170                 175
Leu Ile Val Lys Pro Leu Val Glu Pro Ser Ala Ala Ile Tyr Leu Gln
            180                 185                 190
Gly Leu Thr Asp Leu Lys Thr Ile Glu Leu Lys Ile Asn Thr Ser Pro
        195                 200                 205
Met Lys Ser Leu Ala Phe Leu Lys Pro Leu Phe Gln Arg Gln Ser Gln
    210                 215                 220
Lys Asn Leu Lys Thr Trp Ile Phe Asp Lys Ile Gln Phe Ala Ser Phe
```

```
225                 230                 235                 240
Lys Ile Asp Asn Ala Leu Ile Lys Ala Asn Phe Thr Pro Ser Glu Phe
                245                 250                 255
Ile Pro Ser Leu Leu Glu Asn Ser Val Val Lys Ala Thr Leu Ile Lys
            260                 265                 270
Pro Ser Val Val Phe Asn Asp Gly Leu Ser Pro Ile Lys Met Asp Lys
        275                 280                 285
Thr Glu Leu Ile Phe Lys Asn Lys Gln Leu Leu Ile Gln Pro Gln Lys
    290                 295                 300
Ile Thr Tyr Glu Thr Met Glu Leu Thr Gly Ser Tyr Ala Thr Phe Ser
305                 310                 315                 320
Asn Leu Leu Glu Ala Pro Lys Leu Glu Val Phe Leu Lys Thr Thr Pro
                325                 330                 335
Asn Tyr Tyr Gly Asp Ser Ile Lys Asp Leu Leu Ser Ala Tyr Lys Val
            340                 345                 350
Val Leu Pro Leu Asp Lys Ile Ser Met Pro Ser Ser Ala Asp Leu Lys
        355                 360                 365
Leu Thr Leu Gln Phe Leu Lys Asn Thr Ala Pro Leu Phe Ser Val Gln
    370                 375                 380
Gly Ser Val Asn Leu Gln Glu Gly Thr Phe Ser Leu Tyr Asn Ile Pro
385                 390                 395                 400
Leu Tyr Thr Gln Ser Ala Gln Ile Asn Leu Asp Ile Ala Gln Glu Tyr
                405                 410                 415
Gln Tyr Ile Tyr Ile Asp Thr Ile His Thr Arg Tyr Ala Asn Met Leu
            420                 425                 430
Asp Leu Asp Ala Lys Ile Ala Leu Asp Leu Gly Gln Lys Asn Leu Ser
        435                 440                 445
Leu Asp Ser Leu Val His Lys Ile Gln Val Asn Thr Asn Asn Asn Ile
    450                 455                 460
Asn Met Arg Ser Tyr Asp Pro Asn Asn Thr Gln Glu Asp Pro Gln Thr
465                 470                 475                 480
Asn Phe Thr Leu Asp Leu Lys Ser Leu His Ser Ile Ile Gln Glu Gly
                485                 490                 495
Glu Asn Ser Glu Val Phe Arg Arg Lys Ile Ile Asp Thr Ile Lys Ala
            500                 505                 510
Gln Ser Glu Asp Lys Phe Thr Lys Asp Val Phe Tyr Ala Thr Gly Asp
        515                 520                 525
Thr Leu Lys Ser Leu Ser Leu Ser Phe Asp Phe Ser Asn Pro Asp His
    530                 535                 540
Ile Gln Trp Ser Val Pro Gln Leu Leu Leu Glu Gly Glu Phe Lys Asp
545                 550                 555                 560
Asn Ala Tyr Thr Phe Lys Ile Lys Asp Leu Lys Lys Ile Lys Pro Tyr
                565                 570                 575
Ser Pro Ile Met Asp Tyr Ile Ala Leu Lys Asp Gly Ser Leu Glu Val
            580                 585                 590
Ser Thr Ser Asp Phe Val Asn Ile Asp Phe Phe Ala Lys Asp Leu Lys
        595                 600                 605
Ile Asn Leu Pro Ile Tyr Arg Ser Asp Gly Ser His Phe Asp Ser Phe
    610                 615                 620
Ser Leu Phe Gly Ser Ile Asn Lys Asp Glu Ile Ser Val Tyr Thr Pro
625                 630                 635                 640
Ser Lys Ser Ile Ser Ile Lys Val Lys Gly Asp Gln Lys Asp Ile Thr
                645                 650                 655
Leu Asn Asn Ile Asp Leu Ser Ile Asp Asp Phe Leu Asp Ser Lys Met
            660                 665                 670
```

```
Pro Ala Ile Ala Gly Leu Phe Ser Lys Glu Arg Lys Glu Lys Pro Ser
        675                 680                 685
Ser Lys Glu Ile Gln Asp Glu Asp Val Phe Ile Ser Ala Lys Gln Arg
    690                 695                 700
Tyr Glu Lys Ala His Lys Ile Ile Pro Ile Ser Thr Arg Ile His Ala
705                 710                 715                 720
Lys Asp Val Val Leu Ile Tyr Lys Lys Met Pro Phe Pro Leu Glu Asn
                725                 730                 735
Leu Asp Ile Val Ala Gln Asp Asp Arg Val Lys Ile Asp Gly Asn Tyr
            740                 745                 750
Lys Asn Ala Met Ile Met Ala Asp Leu Val His Gly Ala Leu Tyr Leu
        755                 760                 765
Lys Ala His Asn Phe Ser Gly Asp Tyr Ile Asn Thr Ile Leu Gln Lys
    770                 775                 780
Asp Phe Val Glu Gly Gly Leu Phe Thr Leu Ile Gly Ala Leu Glu Asp
785                 790                 795                 800
Gln Val Phe Asn Gly Glu Leu Lys Phe Gln Asn Thr Ser Leu Lys Asn
                805                 810                 815
Phe Ala Leu Met Gln Asn Met Val Asn Leu Ile Asn Thr Ile Pro Ser
            820                 825                 830
Leu Ile Val Phe Arg Asn Pro His Leu Gly Ala Asn Gly Tyr Gln Ile
        835                 840                 845
Lys Thr Gly Ser Val Val Phe Gly Ile Thr Lys Glu Tyr Leu Gly Leu
    850                 855                 860
Glu Lys Ile Asp Leu Val Gly Lys Thr Leu Asp Ile Ala Gly Asn Gly
865                 870                 875                 880
Ile Ile Glu Leu Asp Lys Asn Lys Leu Asp Leu Asn Leu Glu Val Ser
                885                 890                 895
Thr Ile Lys Ala Leu Ser Asn Val Leu Asn Lys Ile Pro Ile Val Gly
            900                 905                 910
Tyr Leu Val Leu Gly Lys Gly Gly Lys Ile Thr Thr Asn Val Asn Val
        915                 920                 925
Lys Gly Thr Leu Asp Lys Pro Lys Thr Gln Val Thr Leu Ala Ser Asp
    930                 935                 940
Ile Ile Gln Ala Pro Phe Lys Ile Leu Arg Arg Ile Phe Thr Pro Ile
945                 950                 955                 960
Asp Ile Ile Val Asp Glu Val Lys Lys Asn Ile Asp Ser Lys Arg Lys
                965                 970                 975
Leu Lys
```

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 477 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

```
Met Asn Thr Ile Ile Arg Tyr Ala Ser Leu Trp Gly Leu Cys Ile Thr
1               5                  10                  15
Leu Thr Leu Ala Gln Thr Pro Ser Lys Thr Pro Asp Glu Ile Lys Gln
            20                  25                  30
Ile Leu Asn Asn Tyr Ser His Lys Asn Leu Lys Leu Ile Asp Pro Pro
        35                  40                  45
Thr Ser Ser Leu Glu Ala Thr Pro Gly Phe Leu Pro Ser Pro Lys Glu
    50                  55                  60
Thr Ala Thr Thr Ile Asn Gln Glu Ile Ala Lys Tyr His Glu Lys Ser
65                  70                  75                  80
Asp Lys Ala Ala Leu Gly Leu Tyr Glu Leu Leu Lys Gly Ala Thr Thr
                85                  90                  95
Asn Leu Ser Leu Gln Ala Gln Glu Leu Ser Val Lys Gln Ala Met Lys
            100                 105                 110
Asn His Thr Ile Ala Lys Ala Met Phe Leu Pro Thr Leu Asn Ala Ser
        115                 120                 125
Tyr Asn Phe Lys Asn Glu Ala Arg Asp Thr Pro Glu Tyr Lys His Tyr
    130                 135                 140
Asn Thr Gln Gln Leu Gln Ala Gln Val Thr Leu Asn Val Phe Asn Gly
145                 150                 155                 160
Phe Ser Asn Val Asn Asn Val Lys Glu Lys Ser Ala Thr Tyr Arg Ser
                165                 170                 175
Thr Val Ala Asn Leu Glu Tyr Ser Arg Gln Ser Val Tyr Leu Gln Val
            180                 185                 190
Val Gln Gln Tyr Tyr Glu Tyr Phe Asn Asn Leu Ala Arg Met Ile Ala
        195                 200                 205
Leu Gln Lys Lys Leu Glu Gln Ile Gln Thr Asp Ile Lys Arg Val Thr
    210                 215                 220
Lys Leu Tyr Asp Lys Gly Leu Thr Thr Ile Asp Asp Leu Gln Ser Leu
225                 230                 235                 240
Lys Ala Gln Gly Asn Leu Ser Glu Tyr Asp Ile Leu Asp Met Gln Phe
                245                 250                 255
Ala Leu Glu Gln Asn Arg Leu Thr Leu Glu Tyr Leu Thr Asn Leu Ser
            260                 265                 270
Val Lys Asn Leu Lys Lys Thr Thr Ile Asp Ala Pro Asn Leu Gln Leu
        275                 280                 285
Arg Glu Arg Gln Asp Leu Val Ser Leu Arg Glu Gln Ile Ser Ala Leu
    290                 295                 300
Arg Tyr Gln Asn Lys Gln Leu Asn Tyr Tyr Pro Lys Ile Asp Val Phe
305                 310                 315                 320
Asp Ser Trp Leu Phe Trp Ile Gln Lys Pro Ala Tyr Ala Thr Gly Arg
                325                 330                 335
Phe Gly Asn Phe Tyr Pro Gly Gln Gln Asn Thr Ala Gly Val Thr Ala
            340                 345                 350
Thr Leu Asn Ile Phe Asp Asp Ile Gly Leu Ser Leu Gln Lys Gln Ser
        355                 360                 365
Ile Met Leu Gly Gln Leu Ala Asn Glu Lys Asn Leu Ala Tyr Lys Lys
    370                 375                 380
Leu Glu Gln Glu Lys Asp Glu Gln Leu Tyr Arg Lys Ser Leu Asp Ile
```

```
385                 390                 395                 400
Ala Arg Ala Lys Ile Glu Ser Ser Lys Ala Ser Leu Asp Ala Ala Asn
                405                 410                 415
Leu Ser Phe Ala Asn Ile Lys Arg Lys Tyr Asp Ala Asn Leu Val Asp
            420                 425                 430
Phe Thr Thr Tyr Leu Arg Gly Leu Thr Thr Arg Phe Asp Ala Glu Val
        435                 440                 445
Ala Tyr Asn Leu Ala Leu Asn Asn Tyr Glu Val Gln Lys Ala Asn Tyr
    450                 455                 460
Ile Phe Asn Ser Gly His Lys Ile Asp Asp Tyr Val His
465                 470                 475
```

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 412 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...412

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

```
Met Leu Ser Phe Ile Ser Ala Phe Asp Lys Arg Gly Val Ser Ile Arg
1               5                   10                  15
Leu Leu Thr Ala Leu Leu Leu Leu Phe Ser Leu Gly Leu Ala Lys Asp
            20                  25                  30
Leu Glu Ile Gln Thr Phe Val Ala Lys Tyr Leu Ser Lys Asn Gln Lys
        35                  40                  45
Ile Gln Ala Leu Gln Glu Gln Ile Asp Ala Leu Asp Ser Gln Glu Lys
    50                  55                  60
Val Val Ser Lys Trp Asp Asn Pro Ile Leu Tyr Leu Gly Tyr Asn Asn
65                  70                  75                  80
Ala Asn Val Ser Asp Phe Phe Arg Leu Asp Ser Thr Leu Met Gln Asn
                85                  90                  95
Met Ser Leu Gly Leu Ser Gln Lys Val Asp Leu Asn Gly Lys Lys Leu
            100                 105                 110
Thr Gln Ser Lys Met Ile Asn Leu Glu Lys Gln Lys Lys Ile Leu Glu
        115                 120                 125
Leu Lys Lys Thr Lys Gln Gln Leu Val Ile Asn Leu Met Ile Asn Gly
    130                 135                 140
Ile Glu Asn Tyr Lys Asn Gln Gln Glu Ile Glu Leu Leu Asn Thr Ala
145                 150                 155                 160
Ile Lys Asn Leu Glu Asn Thr Leu Tyr Gln Ala Asn His Ser Ser Ser
                165                 170                 175
Pro Asp Leu Ile Ala Ile Ala Lys Leu Glu Ile Leu Lys Ser Leu Leu
```

```
                180                 185                 190
Glu Ile Gln Lys Asn Asp Leu Glu Val Ala Leu Ser Ser Ser His Tyr
        195                 200                 205
Ser Met Gly Glu Leu Thr Phe Lys Glu Asn Glu Ile Leu Ser Ile Ala
    210                 215                 220
Pro Lys Asn Phe Glu Phe Asn Asn Glu Gln Glu Leu His Asn Ile Ser
225                 230                 235                 240
Ala Thr Asn Tyr Asp Ile Ala Ile Ala Arg Leu Asp Glu Glu Lys Ala
                245                 250                 255
Gln Lys Asp Ile Thr Leu Ala Lys Lys Ser Phe Leu Glu Asp Ile Asn
            260                 265                 270
Val Thr Gly Val Tyr Tyr Phe Arg Ser Lys Gln Tyr Tyr Asn Tyr Asp
        275                 280                 285
Met Phe Ser Val Ala Leu Ser Ile Pro Leu Pro Leu Tyr Gly Lys Gln
    290                 295                 300
Ala Lys Leu Val Glu Gln Lys Lys Lys Glu Ser Leu Ala Phe Lys Ser
305                 310                 315                 320
Glu Val Glu Asn Ala Lys Asn Lys Thr Arg His Leu Ala Leu Lys Leu
                325                 330                 335
Leu Lys Lys Leu Glu Thr Leu Gln Lys Asn Leu Glu Ser Ile Asn Lys
            340                 345                 350
Ile Ile Lys Gln Asn Glu Lys Ile Ala Gln Ile Tyr Ala Leu Asp Leu
        355                 360                 365
Lys Thr Asn Gly Asp Tyr Asn Ala Tyr Tyr Asn Ala Leu Asn Asp Lys
    370                 375                 380
Ile Thr Ile Gln Ile Thr Gln Leu Glu Thr Leu Ser Ala Leu Asn Ser
385                 390                 395                 400
Ala Tyr Leu Ser Leu Gln Asn Leu Lys Gly Leu Glu
                405                 410
```

(2) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 137 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...137

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

```
Met Arg Ile Val Arg Asn Leu Phe Leu Val Ser Phe Val Ala Tyr Ser
1               5                   10                  15
Ser Ala Phe Ala Ala Asp Leu Glu Thr Gly Thr Lys Asn Asp Lys Lys
            20                  25                  30
Ser Gly Lys Lys Phe Tyr Lys Leu His Lys Asn His Gly Ser Glu Thr
```

```
          35                  40                  45
Glu Thr Lys Asn Asp Lys Lys Leu Tyr Asp Phe Thr Lys Asn Ser Gly
      50                  55                  60
Leu Glu Gly Val Asp Leu Glu Lys Ser Pro Asn Leu Lys Ser His Lys
65                  70                  75                  80
Lys Ser Asp Lys Lys Phe Tyr Lys Gln Leu Ala Lys Asn Asn Ile Ala
                  85                  90                  95
Glu Gly Val Ser Met Pro Ile Val Asn Phe Asn Lys Ala Leu Ser Phe
             100                 105                 110
Gly Pro Tyr Phe Glu Arg Thr Lys Ser Lys Lys Thr Gln Tyr Met Asp
         115                 120                 125
Gly Gly Leu Met Met His Ile Arg Phe
     130                 135
```

(2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 309 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...309

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

```
Leu Met Pro Gln Asn Gln Leu Val Ile Thr Ile Ile Asp Glu Ser Gly
1               5                   10                  15
Ser Lys Gln Leu Lys Phe Ser Lys Asn Leu Lys Arg Asn Leu Ile Ile
            20                  25                  30
Ser Val Val Ile Leu Leu Leu Ile Val Gly Leu Gly Val Gly Phe Leu
        35                  40                  45
Lys Phe Leu Ile Ala Lys Met Asp Thr Met Thr Ser Glu Arg Asn Ala
    50                  55                  60
Val Leu Arg Asp Phe Arg Gly Leu Tyr Gln Lys Asn Tyr Ala Leu Ala
65                  70                  75                  80
Lys Glu Ile Lys Asn Lys Arg Glu Glu Leu Phe Ile Val Gly Gln Lys
                85                  90                  95
Ile Arg Gly Leu Glu Ser Leu Ile Glu Ile Lys Lys Gly Ala Asn Gly
            100                 105                 110
Gly Gly His Leu Tyr Asp Glu Val Asp Leu Glu Asn Leu Ser Leu Asn
        115                 120                 125
Gln Lys His Leu Ala Leu Met Leu Ile Pro Asn Gly Met Pro Leu Lys
    130                 135                 140
Thr Tyr Ser Ala Ile Lys Pro Thr Lys Glu Arg Asn His Pro Ile Lys
145                 150                 155                 160
Lys Ile Lys Gly Val Glu Ser Gly Ile Asp Phe Ile Ala Pro Leu Asn
```

```
                    165                 170                 175
Thr Pro Val Tyr Ala Ser Ala Asp Gly Ile Val Asp Phe Val Lys Thr
                180                 185                 190
Arg Ser Asn Ala Gly Tyr Gly Asn Leu Val Arg Ile Glu His Ala Phe
        195                 200                 205
Gly Phe Ser Ser Ile Tyr Thr His Leu Asp His Val Asn Val Gln Pro
    210                 215                 220
Lys Ser Phe Ile Gln Lys Gly Gln Leu Ile Gly Tyr Ser Gly Lys Ser
225                 230                 235                 240
Gly Asn Ser Gly Gly Glu Lys Leu His Tyr Glu Val Arg Phe Leu Gly
                245                 250                 255
Lys Ile Leu Asp Ala Glu Lys Phe Leu Ala Trp Asp Leu Asp His Phe
            260                 265                 270
Gln Ser Ala Leu Glu Glu Asn Lys Phe Ile Glu Trp Lys Asn Leu Phe
        275                 280                 285
Trp Val Leu Glu Asp Ile Val Gln Leu Gln Glu His Val Asp Lys Asp
    290                 295                 300
Thr Leu Lys Gly Gln
305
```

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 332 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...332

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

```
Val Leu Tyr Phe Leu Thr Ser Leu Phe Ile Cys Ser Leu Ile Val Leu
1               5                   10                  15
Trp Ser Lys Lys Ser Met Leu Phe Val Asp Asn Ala Asn Lys Ile Gln
            20                  25                  30
Gly Phe His His Ala Arg Thr Pro Arg Ala Gly Gly Leu Gly Ile Phe
        35                  40                  45
Leu Ser Phe Ala Leu Ala Cys Tyr Leu Glu Pro Phe Glu Met Pro Phe
    50                  55                  60
Lys Gly Pro Phe Val Phe Leu Gly Leu Ser Leu Val Phe Leu Ser Gly
65                  70                  75                  80
Phe Leu Glu Asp Ile Asn Leu Ser Leu Ser Pro Lys Ile Arg Leu Ile
                85                  90                  95
Leu Gln Ala Val Gly Val Val Cys Ile Ile Ser Ser Thr Pro Leu Val
            100                 105                 110
Val Ser Asp Phe Ser Pro Leu Phe Ser Leu Pro Tyr Phe Ile Ala Phe
```

```
          115                 120                 125
Leu Phe Ala Ile Phe Met Leu Val Gly Ile Ser Asn Ala Ile Asn Ile
      130                 135                 140
Ile Asp Gly Phe Asn Gly Leu Ala Ser Gly Ile Cys Ala Ile Ala Leu
145                 150                 155                 160
Leu Val Ile His Tyr Ile Asp Pro Ser Ser Leu Ser Cys Leu Leu Ala
                165                 170                 175
Tyr Met Val Leu Gly Phe Met Val Leu Asn Phe Pro Ser Gly Lys Ile
            180                 185                 190
Phe Leu Gly Asp Gly Gly Ala Tyr Phe Leu Gly Leu Val Cys Gly Ile
          195                 200                 205
Ser Leu Leu His Leu Ser Leu Glu Gln Lys Ile Ser Val Phe Phe Gly
      210                 215                 220
Leu Asn Leu Met Leu Tyr Pro Val Ile Glu Val Leu Phe Ser Ile Leu
225                 230                 235                 240
Arg Arg Lys Ile Lys Arg Gln Lys Ala Thr Met Pro Asp Asn Leu His
                245                 250                 255
Leu His Thr Leu Leu Phe Lys Phe Leu Gln Gln Arg Ser Phe Asn Tyr
            260                 265                 270
Pro Asn Pro Leu Cys Ala Phe Ile Leu Ile Leu Cys Asn Leu Pro Phe
          275                 280                 285
Ile Leu Ile Ser Val Leu Phe Arg Leu Asp Ala Tyr Ala Leu Ile Val
      290                 295                 300
Ile Ser Leu Val Phe Ile Ala Cys Tyr Leu Ile Gly Tyr Ala Tyr Leu
305                 310                 315                 320
Asn Arg Gln Val Cys Ala Leu Glu Lys Arg Ala Phe
                325                 330
```

(2) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...271

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

```
Met Asn Ile Phe Lys Arg Ile Ile Cys Val Thr Ala Ile Val Leu Gly
1                 5                  10                  15
Phe Phe Asn Leu Leu Asp Ala Lys His His Lys Glu Lys Lys Glu Asp
            20                  25                  30
His Lys Ile Thr Arg Glu Leu Lys Val Gly Ala Asn Pro Val Pro His
          35                  40                  45
Ala Gln Ile Leu Gln Ser Val Val Asp Asp Leu Lys Glu Lys Gly Ile
```

```
     50                  55                  60
Lys Leu Val Ile Val Ser Phe Thr Asp Tyr Val Leu Pro Asn Leu Ala
65                  70                  75                  80
Leu Asn Asp Gly Ser Leu Asp Ala Asn Tyr Phe Gln His Arg Pro Tyr
                85                  90                  95
Leu Asp Arg Phe Asn Leu Asp Arg Lys Met His Leu Val Gly Leu Ala
            100                 105                 110
Asn Ile His Val Glu Pro Leu Arg Phe Tyr Ser Gln Lys Ile Thr Asp
        115                 120                 125
Ile Lys Asn Leu Lys Lys Gly Ser Val Ile Ala Val Pro Asn Asp Pro
    130                 135                 140
Ala Asn Gln Gly Arg Ala Leu Ile Leu Leu His Lys Gln Gly Leu Ile
145                 150                 155                 160
Ala Leu Lys Asp Pro Ser Asn Leu Tyr Ala Thr Glu Phe Asp Ile Val
                165                 170                 175
Lys Asn Pro Tyr Asn Ile Lys Ile Lys Pro Leu Glu Ala Ala Leu Leu
            180                 185                 190
Pro Lys Val Leu Gly Asp Val Asp Gly Ala Ile Ile Thr Gly Asn Tyr
        195                 200                 205
Ala Leu Gln Ala Lys Leu Thr Gly Ala Leu Phe Ser Glu Asp Lys Asp
    210                 215                 220
Ser Pro Tyr Ala Asn Leu Val Ala Ser Arg Glu Asp Asn Ala Gln Asp
225                 230                 235                 240
Glu Ala Ile Lys Ala Leu Ile Glu Ala Leu Gln Ser Glu Lys Thr Arg
                245                 250                 255
Lys Phe Ile Leu Asp Thr Tyr Lys Gly Ala Ile Ile Pro Ala Phe
            260                 265                 270
```

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 316 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...316

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

```
Met Gln Glu Phe Ser Leu Trp Cys Asp Phe Ile Glu Arg Asp Phe Leu
1               5                   10                  15
Glu Asn Asp Phe Leu Lys Leu Ile Asn Lys Gly Ala Ile Cys Gly Ala
            20                  25                  30
Thr Ser Asn Pro Ser Leu Phe Cys Glu Ala Ile Thr Lys Ser Ala Phe
        35                  40                  45
Tyr Gln Asp Glu Ile Ala Lys Leu Lys Gly Lys Lys Ala Lys Glu Ile
```

```
    50                  55                  60
Tyr Glu Thr Leu Ala Leu Lys Asp Ile Leu Gln Ala Ser Ser Ala Leu
65                  70                  75                  80
Met Pro Leu Tyr Glu Lys Asp Pro Asn Asn Gly Tyr Ile Ser Leu Glu
                85                  90                  95
Ile Asp Pro Phe Leu Glu Asp Asp Ala Ile Lys Ser Ile Asp Glu Ala
            100                 105                 110
Lys Arg Leu Phe Lys Thr Leu Asn Arg Pro Asn Val Met Ile Lys Val
        115                 120                 125
Pro Ala Ser Glu Ser Ala Phe Glu Val Ile Ser Ala Leu Ala Gln Ala
    130                 135                 140
Ser Ile Pro Ile Asn Val Thr Leu Val Phe Ser Pro Lys Ile Ala Gly
145                 150                 155                 160
Glu Ile Ala Gln Ile Leu Ala Lys Glu Ala Arg Lys Arg Ala Val Ile
                165                 170                 175
Ser Val Phe Val Ser Arg Phe Asp Lys Glu Ile Asp Pro Leu Val Pro
            180                 185                 190
Gln Asn Leu Gln Ala Gln Ser Gly Ile Met Asn Ala Thr Glu Cys Tyr
        195                 200                 205
Tyr Gln Ile Asn Gln His Ala Asn Lys Leu Ile Ser Thr Leu Phe Ala
    210                 215                 220
Ser Thr Gly Val Lys Ser Asn Ser Leu Ala Lys Asp Tyr Tyr Ile Lys
225                 230                 235                 240
Ala Leu Cys Phe Lys Asn Ser Ile Asn Thr Ala Pro Leu Asp Ala Leu
                245                 250                 255
Asn Ala Tyr Leu Leu Asp Pro Asn Thr Glu Cys Gln Thr Pro Leu Lys
            260                 265                 270
Ile Thr Glu Ile Glu Ala Phe Lys Lys Glu Leu Lys Thr His Asn Ile
        275                 280                 285
Asp Leu Glu Asn Thr Ala Gln Lys Leu Leu Lys Glu Gly Leu Ile Ala
    290                 295                 300
Phe Lys Gln Ser Phe Glu Lys Leu Leu Ser Ser Phe
305                 310                 315
```

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 260 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...260

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

```
Met Lys Thr Asn Gly His Phe Lys Asp Phe Ala Trp Lys Lys Cys Phe
```

```
1               5                   10                  15
Leu Gly Ala Ser Val Val Ala Leu Leu Val Gly Cys Ser Pro His Ile
            20                  25                  30
Ile Glu Thr Asn Glu Val Ala Leu Lys Leu Asn Tyr His Pro Ala Ser
        35                  40                  45
Glu Lys Val Gln Ala Leu Asp Glu Lys Ile Leu Leu Leu Arg Pro Ala
    50                  55                  60
Phe Gln Tyr Ser Asp Asn Ile Ala Lys Glu Tyr Glu Asn Lys Phe Lys
65                  70                  75                  80
Asn Gln Thr Thr Leu Lys Val Glu Glu Ile Leu Gln Asn Gln Gly Tyr
                85                  90                  95
Lys Val Ile Asn Val Asp Ser Ser Asp Lys Asp Asp Phe Ser Phe Ala
            100                 105                 110
Gln Lys Lys Glu Gly Tyr Leu Ala Val Ala Met Asn Gly Glu Ile Val
        115                 120                 125
Leu Arg Pro Asp Pro Lys Arg Thr Ile Gln Lys Lys Ser Glu Pro Gly
    130                 135                 140
Leu Leu Phe Ser Thr Gly Leu Asp Lys Met Glu Arg Val Leu Ile Pro
145                 150                 155                 160
Ala Gly Phe Val Lys Val Thr Ile Leu Glu Pro Met Ser Gly Glu Ser
                165                 170                 175
Leu Asp Ser Phe Thr Met Asp Leu Ser Glu Leu Asp Ile Gln Glu Lys
            180                 185                 190
Phe Leu Lys Thr Thr His Ser Ser His Ser Gly Gly Leu Val Ser Thr
        195                 200                 205
Met Val Lys Gly Thr Asp Asn Ser Asn Asp Ala Ile Lys Ser Ala Leu
    210                 215                 220
Asn Lys Ile Phe Ala Ser Ile Met Gln Glu Met Asp Lys Lys Leu Thr
225                 230                 235                 240
Gln Arg Asn Leu Glu Ser Tyr Gln Lys Asp Ala Lys Glu Leu Lys Asn
                245                 250                 255
Lys Arg Asn Arg
            260
```

(2) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1382 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1382

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

Leu Asn Phe Asn Asn Leu Thr Ala Asn Gly Ala Leu Asn Phe Asn Gly

```
1               5                  10                  15
Tyr Ala Pro Ser Leu Thr Lys Ala Leu Met Asn Val Ser Gly Gln Phe
            20                  25                  30
Val Leu Gly Asn Asn Gly Asp Ile Asn Leu Ser Asp Ile Asn Ile Phe
        35                  40                  45
Asp Asn Ile Thr Lys Ser Val Thr Tyr Asn Ile Leu Asn Ala Gln Lys
    50                  55                  60
Gly Ile Thr Gly Ile Ser Gly Ala Asn Gly Tyr Glu Lys Ile Leu Phe
65                  70                  75                  80
Tyr Gly Met Lys Ile Gln Asn Ala Thr Tyr Ser Asp Asn Asn Asn Ile
                85                  90                  95
Gln Thr Trp Ser Phe Ile Asn Pro Leu Asn Ser Ser Gln Ile Ile Gln
            100                 105                 110
Glu Ser Ile Lys Asn Gly Asp Leu Thr Ile Glu Val Leu Asn Asn Pro
        115                 120                 125
Asn Ser Ala Ser Asn Thr Ile Phe Asn Ile Ala Pro Glu Leu Tyr Asn
    130                 135                 140
Tyr Gln Asp Ser Lys Gln Asn Pro Thr Gly Tyr Ser Tyr Asp Tyr Ser
145                 150                 155                 160
Asp Asn Gln Ala Gly Thr Tyr Tyr Leu Thr Ser Asn Ile Lys Gly Leu
                165                 170                 175
Phe Thr Pro Lys Gly Ser Gln Thr Pro Gln Thr Pro Gly Thr Tyr Ser
            180                 185                 190
Pro Phe Asn Gln Pro Leu Asn Ser Leu Asn Ile Tyr Asn Lys Gly Phe
        195                 200                 205
Ser Ser Glu Asn Leu Lys Thr Leu Leu Gly Ile Leu Ser Gln Asn Ser
    210                 215                 220
Ala Thr Leu Lys Glu Met Ile Glu Ser Asn Gln Leu Asp Asn Ile Thr
225                 230                 235                 240
Asn Ile Asn Glu Val Leu Gln Leu Leu Asp Lys Ile Lys Ile Thr Gln
                245                 250                 255
Ala Gln Lys Gln Ala Leu Leu Glu Thr Ile Asn His Leu Thr Asp Asn
            260                 265                 270
Ile Asn Gln Thr Phe Asn Asn Gly Asn Leu Val Ile Gly Ala Thr Gln
        275                 280                 285
Asp Asn Val Thr Asn Ser Thr Ser Ser Ile Trp Phe Gly Gly Asn Gly
    290                 295                 300
Tyr Ser Ser Pro Cys Ala Leu Asp Ser Ala Thr Cys Ser Ser Phe Arg
305                 310                 315                 320
Asn Thr Tyr Leu Gly Gln Leu Leu Gly Ser Thr Ser Pro Tyr Leu Gly
                325                 330                 335
Tyr Ile Asn Ala Asp Phe Lys Ala Lys Ser Ile Tyr Ile Thr Gly Thr
            340                 345                 350
Ile Gly Ser Ser Asn Ala Phe Glu Ser Gly Gly Ser Ala Asp Val Thr
        355                 360                 365
Phe Gln Ser Ala Asn Asn Leu Val Leu Asn Lys Ala Asn Ile Glu Ala
    370                 375                 380
Gln Ala Thr Asp Asn Ile Phe Asn Leu Leu Gly Gln Glu Gly Ile Asp
385                 390                 395                 400
Lys Ile Phe Asn Gln Gly Asn Leu Ala Asn Val Leu Ser Gln Met Ala
                405                 410                 415
Met Glu Lys Ile Lys Gln Ala Gly Gly Leu Gly Asn Phe Ile Glu Asn
            420                 425                 430
Ala Leu Ser Pro Leu Ser Lys Glu Leu Pro Ala Ser Leu Gln Asp Glu
        435                 440                 445
```

Thr Leu Gly Gln Leu Ile Gly Gln Asn Asn Leu Asp Asp Leu Leu Asn
450 455 460
Asn Ser Gly Val Met Asn Glu Ile Gln Asn Ile Ile Ser Gln Lys Leu
465 470 475 480
Ser Ile Phe Gly Asn Phe Val Thr Pro Ser Ile Ile Glu Asn Tyr Leu
485 490 495
Ala Lys Gln Ser Leu Lys Ser Met Leu Asp Asp Lys Gly Leu Leu Asn
500 505 510
Phe Ile Gly Gly Tyr Ile Asp Ala Ser Glu Leu Ser Ser Ile Leu Gly
515 520 525
Val Ile Leu Lys Asp Ile Thr Asn Pro Pro Thr Ser Leu Gln Lys Asp
530 535 540
Ile Gly Val Val Ala Asn Asp Leu Leu Asn Glu Phe Leu Gly Gln Asp
545 550 555 560
Val Val Lys Lys Leu Glu Ser Gln Gly Leu Val Ser Asn Ile Ile Asn
565 570 575
Asn Val Ile Ser Gln Gly Gly Leu Ser Gly Val Tyr Asn Gln Gly Leu
580 585 590
Gly Ser Val Leu Pro Pro Ser Leu Gln Asn Ala Leu Lys Glu Asn Asp
595 600 605
Leu Gly Thr Leu Leu Ser Pro Arg Gly Leu His Asp Phe Trp Gln Lys
610 615 620
Gly Tyr Phe Asn Phe Leu Ser Asn Gly Tyr Val Phe Val Asn Asn Ser
625 630 635 640
Ser Phe Ser Asn Ala Thr Gly Gly Ser Leu Asn Phe Val Ala Asn Lys
645 650 655
Ser Ile Ile Phe Asn Gly Asp Asn Thr Ile Asp Phe Ser Lys Tyr Gln
660 665 670
Gly Ala Leu Ile Phe Ala Ser Asn Gly Val Ser Asn Ile Asn Ile Thr
675 680 685
Thr Leu Asn Ala Thr Asn Gly Leu Ser Leu Asn Ala Gly Leu Asn Asn
690 695 700
Val Ser Val Gln Lys Gly Glu Ile Cys Ile Asn Leu Ala Asn Cys Pro
705 710 715 720
Thr Thr Lys Asn Ser Ser Pro Ala Asn Ser Ser Val Thr Pro Thr Asn
725 730 735
Glu Ser Leu Ser Val His Ala Asn Asn Phe Thr Phe Leu Gly Thr Ile
740 745 750
Ile Ser Asn Gly Ala Ile Asp Leu Ser Gln Val Thr Asn Asn Ser Val
755 760 765
Ile Gly Thr Leu Asn Leu Asn Glu Asn Ala Thr Leu Gln Ala Asn Asn
770 775 780
Leu Thr Ile Thr Asn Ala Phe Asn Asn Ala Ser Asn Ser Thr Ala Asn
785 790 795 800
Ile Asp Gly Asn Phe Thr Leu Asn Gln Gln Ala Thr Leu Ser Thr Asn
805 810 815
Ala Ser Gly Leu Asn Val Met Gly Asn Phe Asn Ser Tyr Gly Asp Leu
820 825 830
Val Phe Asn Leu Ser His Ser Val Ser His Ala Ile Ile Asn Thr Gln
835 840 845
Gly Thr Ala Thr Ile Met Ala Asn Asn Asn Pro Leu Ile Gln Phe Asn
850 855 860
Ala Ser Ser Lys Glu Val Gly Thr Tyr Thr Leu Ile Asp Ser Ala Lys
865 870 875 880
Ala Ile Tyr Tyr Gly Tyr Asn Asn Gln Ile Thr Gly Gly Ser Ser Leu

885 890 895
Asp Asn Tyr Leu Lys Leu Tyr Ala Leu Ile Asp Ile Asn Gly Lys His
900 905 910
Met Val Met Thr Asp Asn Gly Leu Thr Tyr Asn Gly Gln Ala Val Ser
915 920 925
Val Lys Asp Gly Gly Leu Val Val Gly Phe Lys Asp Ser Gln Asn Gln
930 935 940
Tyr Ile Tyr Thr Ser Ile Leu Tyr Asn Lys Val Lys Ile Ala Val Ser
945 950 955 960
Asn Asp Pro Ile Asn Asn Pro Gln Ala Pro Thr Leu Lys Gln Tyr Ile
965 970 975
Ala Gln Ile Gln Gly Val Gln Ser Val Asp Ser Ile Asp Gln Ala Gly
980 985 990
Gly Asn Gln Ala Ile Asn Trp Leu Asn Lys Ile Phe Glu Thr Lys Gly
995 1000 1005
Ser Pro Leu Phe Ala Pro Tyr Tyr Leu Glu Ser His Ser Thr Lys Asp
1010 1015 1020
Leu Thr Thr Ile Ala Gly Asp Ile Ala Asn Thr Leu Glu Val Ile Ala
1025 1030 1035 1040
Asn Pro Asn Phe Lys Asn Asp Ala Thr Asn Ile Leu Gln Ile Asn Thr
1045 1050 1055
Tyr Thr Gln Gln Met Ser Arg Leu Ala Lys Leu Ser Asp Thr Ser Thr
1060 1065 1070
Phe Ala Arg Ser Asp Phe Leu Glu Arg Leu Glu Ala Leu Lys Asn Lys
1075 1080 1085
Arg Phe Ala Asp Ala Ile Pro Asn Ala Met Asp Val Ile Leu Lys Tyr
1090 1095 1100
Ser Gln Arg Asn Arg Val Lys Asn Asn Val Trp Ala Thr Gly Val Gly
1105 1110 1115 1120
Gly Ala Ser Phe Ile Ser Gly Gly Thr Gly Thr Leu Tyr Gly Ile Asn
1125 1130 1135
Val Gly Tyr Asp Arg Phe Ile Lys Gly Val Ile Val Gly Gly Tyr Ala
1140 1145 1150
Ala Tyr Gly Tyr Ser Gly Phe His Ala Asn Ile Thr Gln Ser Gly Ser
1155 1160 1165
Ser Asn Val Asn Val Gly Val Tyr Ser Arg Ala Phe Ile Lys Arg Ser
1170 1175 1180
Glu Leu Thr Met Ser Leu Asn Glu Thr Trp Gly Tyr Asn Lys Thr Phe
1185 1190 1195 1200
Ile Asn Ser Tyr Asp Pro Leu Leu Ser Ile Ile Asn Gln Ser Tyr Arg
1205 1210 1215
Tyr Asp Thr Trp Thr Thr Asp Ala Lys Ile Asn Tyr Gly Tyr Asp Phe
1220 1225 1230
Met Phe Lys Asp Lys Ser Val Ile Phe Lys Pro Gln Val Gly Leu Ser
1235 1240 1245
Tyr Tyr Tyr Ile Gly Leu Ser Gly Leu Arg Gly Ile Met Asp Asp Pro
1250 1255 1260
Ile Tyr Asn Gln Phe Arg Ala Asn Ala Asp Pro Asn Lys Lys Ser Val
1265 1270 1275 1280
Leu Thr Ile Asn Phe Ala Leu Glu Ser Arg His Tyr Phe Asn Lys Asn
1285 1290 1295
Ser Tyr Tyr Phe Val Ile Ala Asp Val Gly Arg Asp Leu Phe Ile Asn
1300 1305 1310
Ser Met Gly Asp Lys Met Val Arg Phe Ile Gly Asn Asn Thr Leu Ser
1315 1320 1325

```
Tyr Arg Asp Gly Gly Arg Tyr Asn Thr Phe Ala Ser Ile Ile Thr Gly
    1330                1335                1340
Gly Glu Ile Arg Leu Phe Lys Thr Phe Tyr Val Asn Ala Gly Ile Gly
1345                1350                1355                1360
Ala Arg Phe Gly Leu Asp Tyr Lys Asp Ile Asn Ile Thr Gly Asn Ile
                1365                1370                1375
Gly Met Arg Tyr Ala Phe
            1380
```

(2) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 262 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...262

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

```
Met Lys Lys Ile Gly Leu Ser Leu Cys Leu Val Leu Ser Leu Gly Phe
1               5                   10                  15
Leu Lys Ala His Glu Val Ser Ala Glu Glu Ile Ala Asp Ile Phe Tyr
            20                  25                  30
Lys Leu Asn Ala Lys Glu Pro Lys Met Lys Ile Asn His Thr Lys Gly
        35                  40                  45
Phe Cys Ala Lys Gly Val Phe Leu Pro Asn Pro Gln Ala Arg Glu Asp
    50                  55                  60
Leu Glu Val Pro Leu Leu Asn Glu Lys Glu Ile Pro Ala Ser Val Arg
65                  70                  75                  80
Tyr Ser Leu Gly Gly Val Ala Met Asp Asp Lys Ser Lys Val Arg Gly
                85                  90                  95
Met Ala Leu Lys Leu Glu Asn Gln Asn Ala Ser Trp Thr Met Val Met
            100                 105                 110
Leu Asn Thr Glu Ile Asn Phe Ala Lys Asn Pro Glu Glu Phe Ala Gln
        115                 120                 125
Phe Phe Glu Met Arg Leu Pro Lys Asn Gly Lys Val Asp Glu Ala Arg
    130                 135                 140
Ile Lys Lys Leu Tyr Glu Glu Val Pro Ser Tyr Arg Asn Phe Ala Ala
145                 150                 155                 160
Tyr Met Lys Thr Ile Gly Ile Ser Ser Ser Val Ala Asn Thr Pro Tyr
                165                 170                 175
Tyr Ser Val His Ala Phe Lys Phe Lys Asp Lys Lys Glu Lys Leu Leu
            180                 185                 190
Pro Ala Arg Trp Lys Phe Val Pro Lys Glu Gly Val Lys Tyr Leu Asn
        195                 200                 205
```

```
Pro Gln Glu Leu Lys Gln Lys Asp Ser Asn Tyr Leu Leu Ser Ser Phe
    210                 215                 220
Gln Gln His Leu Lys Asn Lys Pro Ile Glu Tyr Gln Met Tyr Leu Val
225                 230                 235                 240
Phe Ala Asn Gln Asn Asp Ala Thr Asn Asp Thr Thr Ala Leu Trp Lys
                245                 250                 255
Gly Ser Ile Arg Asn Tyr
            260
```

(2) INFORMATION FOR SEQ ID NO:133:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 246 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...246

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

```
Met Lys Gln Phe Lys Lys Lys Pro Lys Lys Ile Lys Arg Ser His Gln
1               5                   10                  15
Asn Gln Lys Thr Ile Leu Lys Arg Pro Leu Trp Leu Met Pro Leu Leu
            20                  25                  30
Ile Gly Gly Phe Ala Ser Gly Val Tyr Ala Asp Gly Thr Asp Ile Leu
        35                  40                  45
Gly Leu Ser Trp Gly Glu Lys Ser Gln Lys Val Cys Val His Arg Pro
    50                  55                  60
Trp Tyr Ala Ile Trp Ser Cys Asp Lys Trp Glu Glu Lys Thr Gln Gln
65                  70                  75                  80
Phe Thr Gly Asn Gln Leu Ile Thr Lys Thr Trp Ala Gly Gly Asn Ala
                85                  90                  95
Ala Asn Tyr Tyr His Ser Gln Asn Asn Gln Asp Ile Thr Ala Asn Leu
            100                 105                 110
Lys Asn Asp Asn Gly Thr Tyr Phe Leu Ser Gly Leu Tyr Asn Tyr Thr
        115                 120                 125
Gly Gly Glu Tyr Asn Gly Gly Asn Leu Asp Ile Glu Leu Gly Ser Asn
    130                 135                 140
Ala Thr Phe Asn Leu Gly Ala Ser Ser Gly Asn Ser Phe Thr Ser Trp
145                 150                 155                 160
Tyr Pro Asn Gly His Thr Asp Val Thr Phe Ser Ala Gly Thr Ile Asn
                165                 170                 175
Val Asn Asn Ser Val Glu Val Gly Asn Arg Val Gly Ser Gly Ala Gly
            180                 185                 190
Thr His Thr Gly Thr Ala Thr Leu Asn Leu Asn Ala Asn Lys Val Thr
        195                 200                 205
```

```
Ile Asn Ser Asn Ile Ser Ala Tyr Lys Thr Ser Gln Val Asn Val Gly
    210                 215                 220
Asn Ala Asn Ser Val Ile Thr Ile Asn Ser Val Ser Leu Asn Gly Glu
225                 230                 235                 240
Tyr Leu Gln Phe Phe Ser
                245
```

(2) INFORMATION FOR SEQ ID NO:134:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...245

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

```
Met Ile Lys Lys Thr Leu Ala Ser Val Leu Leu Gly Leu Ser Leu Met
1               5                   10                  15
Ser Val Leu Asn Ala Lys Glu Cys Val Ser Pro Ile Thr Arg Ser Val
            20                  25                  30
Lys Tyr His Gln Gln Ser Ala Glu Ile Arg Ala Leu Gln Leu Gln Ser
        35                  40                  45
Tyr Lys Met Ala Lys Met Ala Leu Asp Asn Asn Leu Lys Leu Val Lys
    50                  55                  60
Asp Lys Lys Pro Ala Val Ile Leu Asp Leu Asp Glu Thr Val Leu Asn
65                  70                  75                  80
Thr Phe Asp Tyr Ala Gly Tyr Leu Val Lys Asn Cys Ile Lys Tyr Thr
                85                  90                  95
Pro Glu Thr Trp Asp Lys Phe Glu Lys Glu Gly Ser Leu Thr Leu Ile
            100                 105                 110
Pro Gly Ala Leu Asp Phe Leu Glu Tyr Ala Asn Ser Lys Gly Val Lys
        115                 120                 125
Ile Phe Tyr Ile Ser Asn Arg Thr Gln Lys Asn Lys Ala Phe Thr Leu
    130                 135                 140
Lys Thr Leu Lys Ser Phe Lys Leu Pro Gln Val Ser Glu Glu Ser Val
145                 150                 155                 160
Leu Leu Lys Glu Lys Gly Lys Pro Lys Ala Val Arg Arg Glu Leu Val
                165                 170                 175
Ala Lys Asp Tyr Ala Ile Val Leu Gln Val Gly Asp Thr Leu His Asp
            180                 185                 190
Phe Asp Ala Ile Phe Ala Lys Asp Ala Lys Asn Ser Gln Glu Gln Gln
        195                 200                 205
Ala Lys Val Leu Gln Asn Ala Gln Lys Phe Gly Thr Glu Trp Ile Ile
    210                 215                 220
```

```
Leu Pro Asn Ser Leu Tyr Gly Thr Trp Glu Asp Gly Pro Ile Lys Ala
225                 230                 235                 240
Trp Gln Asn Lys Lys
                245
```

(2) INFORMATION FOR SEQ ID NO:135:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 288 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...288

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

```
Leu Trp Cys Leu Lys Thr Pro Ile Ile Gly His Gly Met Lys Lys Lys
1               5                  10                  15
Ala Lys Val Phe Trp Cys Cys Phe Lys Met Ile Arg Trp Leu Tyr Leu
            20                  25                  30
Ala Val Phe Phe Leu Leu Ser Val Ser Asp Ala Lys Glu Ile Ala Met
        35                  40                  45
Gln Arg Phe Asp Lys Gln Asn His Lys Ile Phe Glu Ile Leu Ala Asp
    50                  55                  60
Lys Val Ser Ala Lys Asp Asn Val Ile Thr Ala Ser Gly Asn Ala Ile
65                  70                  75                  80
Leu Leu Asn Tyr Asp Val Tyr Ile Leu Ala Asp Lys Val Arg Tyr Asp
                85                  90                  95
Thr Lys Thr Lys Glu Ala Leu Leu Glu Gly Asn Ile Lys Val Tyr Arg
            100                 105                 110
Gly Glu Gly Leu Leu Val Lys Thr Asp Tyr Val Lys Leu Ser Leu Asn
        115                 120                 125
Glu Lys Tyr Glu Ile Ile Phe Pro Phe Tyr Val Gln Asp Ser Val Ser
    130                 135                 140
Gly Ile Trp Val Ser Ala Asp Ile Ala Ser Gly Lys Asp Gln Lys Tyr
145                 150                 155                 160
Lys Ile Lys Asn Met Ser Ala Ser Gly Cys Ser Ile Asp Asn Pro Ile
                165                 170                 175
Trp His Val Asn Ala Thr Ser Gly Ser Phe Asn Met Gln Lys Ser His
            180                 185                 190
Leu Ser Met Trp Asn Pro Lys Ile Tyr Val Gly Asp Ile Pro Val Leu
        195                 200                 205
Tyr Leu Pro Tyr Ile Phe Met Ser Thr Ser Asn Lys Arg Thr Thr Gly
    210                 215                 220
Phe Leu Tyr Pro Glu Phe Gly Thr Ser Asn Leu Asp Gly Phe Ile Tyr
225                 230                 235                 240
```

```
Leu Gln Pro Phe Tyr Leu Ala Pro Lys Asn Ser Trp Asp Met Thr Phe
                245                 250                 255
Thr Pro Gln Ile Arg Tyr Lys Arg Gly Phe Gly Leu Asn Phe Glu Ala
            260                 265                 270
Arg Tyr Ile Asn Ser Lys Thr Gln Val Phe Ile Gln Cys Ala Leu Phe
        275                 280                 285
```

(2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 128 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...128

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

```
Leu Met Phe Lys Lys Met Cys Leu Ser Leu Leu Met Ile Ser Gly Val
1               5                  10                  15
Cys Val Gly Ala Lys Asp Leu Asp Phe Lys Leu Asp Tyr Arg Ala Thr
            20                  25                  30
Gly Gly Lys Phe Met Gly Lys Met Thr Asp Ser Ser Leu Leu Ser Ile
        35                  40                  45
Thr Ser Met Asn Asp Glu Pro Val Val Ile Lys Asn Leu Ile Val Asn
    50                  55                  60
Arg Gly Asn Ser Cys Glu Ala Thr Lys Lys Val Glu Pro Lys Phe Gly
65                  70                  75                  80
Asp Lys Phe Lys Lys Glu Lys Leu Phe Asp His Glu Leu Lys Tyr Ser
                85                  90                  95
Gln Gln Ile Phe Tyr Arg Leu Asp Cys Lys Pro Asn Gln Leu Leu Glu
            100                 105                 110
Val Lys Ile Ile Thr Asp Lys Gly Glu Tyr Tyr His Lys Phe Ser Lys
        115                 120                 125
```

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 169 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...169

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

```
Met Gln Ala Leu Lys Ser Leu Leu Glu Val Ile Thr Lys Leu Gln Asn
1               5                  10                  15
Leu Gly Gly Tyr Leu Met His Ile Ala Ile Phe Ile Ile Phe Ile Trp
            20                  25                  30
Ile Gly Gly Leu Lys Phe Val Pro Tyr Glu Ala Glu Gly Ile Ala Pro
        35                  40                  45
Phe Val Ala Asn Ser Pro Phe Phe Ser Phe Met Tyr Lys Phe Glu Lys
    50                  55                  60
Pro Ala Tyr Lys Gln His Lys Met Ser Glu Ser Gln Ser Met Gln Glu
65                  70                  75                  80
Glu Met Gln Asp Asn Pro Lys Ile Val Glu Asn Lys Glu Trp His Lys
                85                  90                  95
Glu Asn Arg Thr Tyr Leu Val Ala Glu Gly Leu Gly Ile Thr Ile Met
            100                 105                 110
Ile Leu Gly Ile Leu Val Leu Leu Gly Leu Trp Met Pro Leu Met Gly
        115                 120                 125
Val Val Gly Gly Leu Leu Val Ala Gly Met Thr Ile Thr Thr Leu Phe
    130                 135                 140
Phe Phe Ile His Asn Ala Arg Ser Val Cys Gln Ser Ala Phe Pro Met
145                 150                 155                 160
Ala Phe Trp Gly Trp Lys Ala Ser Gly
                165
```

(2) INFORMATION FOR SEQ ID NO:138:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 487 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...487

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

```
Met Ile Glu Trp Met Gln Asn His Arg Lys Tyr Leu Val Val Thr Ile
1               5                  10                  15
Trp Ile Ser Thr Ile Ala Phe Ile Ala Ala Gly Met Ile Gly Trp Gly
```

```
                20                  25                  30
Gln Tyr Ser Phe Ser Leu Asp Ser Asp Ser Ala Ala Lys Val Gly Gln
        35                  40                  45
Ile Lys Ile Ser Gln Glu Glu Leu Ala Gln Glu Tyr Arg Arg Leu Lys
    50                  55                  60
Asp Ala Tyr Ala Glu Ser Ile Pro Asp Phe Lys Glu Leu Thr Glu Asp
65                  70                  75                  80
Gln Ile Lys Ala Met His Leu Glu Lys Ser Ala Leu Asp Ser Leu Ile
                85                  90                  95
Asn Gln Ala Leu Leu Arg Asn Phe Ala Leu Asp Leu Gly Leu Gly Ala
            100                 105                 110
Thr Lys Gln Glu Val Ala Lys Glu Ile Arg Lys Thr Asn Val Phe Gln
        115                 120                 125
Lys Asp Gly Val Phe Asp Glu Glu Leu Tyr Lys Asn Ile Leu Lys Gln
    130                 135                 140
Ser His Tyr Arg Pro Lys His Phe Glu Glu Ser Val Glu Arg Leu Leu
145                 150                 155                 160
Ile Leu Gln Lys Ile Ser Ala Leu Phe Pro Lys Thr Thr Thr Pro Leu
                165                 170                 175
Glu Gln Ser Ser Leu Ser Leu Trp Ala Lys Leu Gln Asp Lys Leu Asp
            180                 185                 190
Ile Leu Ile Leu Asn Pro Asn Asp Val Lys Ile Ser Leu Asn Glu Glu
        195                 200                 205
Glu Met Lys Lys Tyr Tyr Glu Asn His Arg Lys Asp Phe Lys Lys Pro
    210                 215                 220
Thr Ser Phe Lys Thr Arg Ser Leu Tyr Phe Asp Ala Ser Leu Glu Lys
225                 230                 235                 240
Thr Asp Leu Lys Glu Leu Glu Glu Tyr Tyr His Lys Asn Lys Val Ser
                245                 250                 255
Tyr Leu Asp Lys Glu Gly Lys Leu Gln Asp Phe Lys Ser Val Gln Glu
            260                 265                 270
Gln Val Lys His Asp Leu Asn Met Gln Lys Ala Asn Glu Lys Ala Leu
        275                 280                 285
Arg Ser Tyr Ile Ala Leu Lys Lys Gly Asn Ala Gln Asn Tyr Thr Thr
    290                 295                 300
Gln Asp Phe Glu Lys Asn Asn Ser Pro Tyr Thr Ala Glu Ile Thr Gln
305                 310                 315                 320
Lys Leu Thr Ala Leu Lys Pro Leu Glu Val Leu Lys Pro Glu Pro Phe
                325                 330                 335
Lys Asp Gly Phe Ile Val Val Gln Leu Val Ser Gln Ile Lys Asp Glu
            340                 345                 350
Leu Gln Asn Phe Asp Glu Ala Lys Ser Ala Leu Lys Thr Arg Leu Thr
        355                 360                 365
Gln Glu Lys Thr Leu Met Ala Leu Gln Thr Leu Ala Lys Glu Lys Leu
    370                 375                 380
Lys Asp Phe Lys Gly Lys Ser Val Gly Tyr Val Ser Pro Asn Phe Gly
385                 390                 395                 400
Gly Thr Ile Ser Glu Leu Asn Gln Glu Glu Ser Ala Lys Phe Ile Asn
                405                 410                 415
Thr Leu Phe Asn Arg Gln Glu Lys Lys Gly Phe Val Thr Ile Gly Asn
            420                 425                 430
Lys Val Val Leu Tyr Gln Ile Thr Glu Gln Asn Phe Asn His Pro Phe
        435                 440                 445
Ser Ala Glu Glu Asn Gln Tyr Met Gln Arg Leu Val Asn Asn Thr Lys
    450                 455                 460
```

```
Thr Asp Phe Phe Asp Lys Ala Leu Ile Glu Glu Leu Lys Lys Arg Tyr
465                 470                 475                 480
Lys Ile Val Lys Tyr Ile Gln
                485
```

(2) INFORMATION FOR SEQ ID NO:139:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 142 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...142

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

```
Met Lys Thr Asn Phe Tyr Lys Ile Lys Leu Leu Phe Ala Trp Cys Leu
1               5                  10                  15
Ile Ile Gly Met Phe Asn Ala Pro Leu Asn Ala Asp Gln Asn Thr Asp
            20                  25                  30
Ile Lys Asp Ile Ser Pro Glu Asp Met Ala Leu Asn Ser Val Gly Leu
        35                  40                  45
Val Ser Arg Asp Gln Leu Lys Ile Glu Ile Pro Lys Glu Thr Leu Glu
    50                  55                  60
Gln Lys Val Ala Ile Leu Asn Asp Tyr Asn Asp Lys Asn Val Asn Ile
65                  70                  75                  80
Lys Phe Asp Asp Ile Ser Leu Gly Ser Phe Gln Pro Asn Asp Asn Leu
                85                  90                  95
Gly Ile Asn Ala Met Trp Gly Ile Gln Asn Leu Leu Met Ser Gln Met
            100                 105                 110
Met Ser Asn Tyr Gly Pro Asn Asn Ser Phe Met Tyr Gly Tyr Ala Pro
        115                 120                 125
Thr Tyr Ser Asp Ser Ser Phe Leu Pro Pro Ile Leu Gly Tyr
    130                 135                 140
```

(2) INFORMATION FOR SEQ ID NO:140:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 208 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...208

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

```
Leu Ile Asn Asn Asn Asn Asn Asn Lys Lys Leu Arg Gly Phe Phe Leu
1               5                  10                  15
Lys Val Leu Leu Ser Leu Val Val Phe Ser Ser Tyr Gly Ser Ala Asn
            20                  25                  30
Asp Asp Lys Glu Ala Lys Lys Glu Ala Leu Glu Lys Glu Lys Asn Thr
        35                  40                  45
Pro Asn Gly Leu Val Tyr Thr Asn Leu Asp Phe Asp Ser Phe Lys Ala
    50                  55                  60
Thr Ile Lys Asn Leu Lys Asp Lys Lys Val Thr Phe Lys Glu Val Asn
65                  70                  75                  80
Pro Asp Ile Ile Lys Asp Glu Val Phe Asp Phe Val Ile Val Asn Arg
                85                  90                  95
Val Leu Lys Lys Ile Lys Asp Leu Lys His Tyr Asp Pro Val Ile Glu
            100                 105                 110
Lys Ile Phe Asp Glu Lys Gly Lys Glu Met Gly Leu Asn Val Glu Leu
        115                 120                 125
Gln Ile Asn Pro Glu Val Lys Asp Phe Phe Thr Phe Lys Ser Ile Ser
    130                 135                 140
Thr Thr Asn Lys Gln Arg Cys Phe Leu Ser Leu His Gly Glu Thr Arg
145                 150                 155                 160
Glu Ile Leu Cys Asp Asp Lys Leu Tyr Asn Val Leu Leu Ala Val Phe
                165                 170                 175
Asn Ser Tyr Asp Pro Asn Asp Leu Leu Lys His Ile Ser Thr Ile Glu
            180                 185                 190
Ser Leu Lys Lys Ile Phe Tyr Thr Ile Thr Cys Glu Ala Val Tyr Leu
        195                 200                 205
```

(2) INFORMATION FOR SEQ ID NO:141:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...245

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

```
Met Ala Gly Thr Gln Ala Ile Tyr Glu Ser Ser Ser Ala Gly Phe Leu
1               5                  10                  15
Ser Gln Val Ser Ser Ile Ile Ser Ser Thr Ser Gly Val Ala Gly Pro
            20                  25                  30
Phe Ala Gly Ile Val Ala Gly Ala Met Thr Ala Ala Ile Ile Pro Ile
        35                  40                  45
Val Val Gly Phe Thr Asn Pro Gln Met Thr Ala Ile Met Thr Gln Tyr
    50                  55                  60
Asn Gln Ser Ile Ala Glu Ala Val Ser Val Pro Met Lys Ala Ala Asn
65                  70                  75                  80
Gln Gln Tyr Asn Gln Leu Tyr Gln Gly Phe Asn Asp Gln Ser Met Ala
                85                  90                  95
Val Gly Asn Asn Ile Leu Asn Ile Ser Lys Leu Thr Gly Glu Phe Asn
            100                 105                 110
Ala Gln Gly Asn Thr Gln Ser Ala Gln Ile Ser Ala Val Asn Ser Gln
        115                 120                 125
Ile Ala Ser Ile Leu Ala Ser Asn Thr Thr Pro Lys Asn Pro Ser Ala
    130                 135                 140
Ile Glu Ala Tyr Ala Thr Asn Gln Ile Ala Val Pro Ser Val Pro Thr
145                 150                 155                 160
Thr Val Glu Met Met Ser Gly Ile Leu Gly Asn Ile Thr Ser Ala Ala
                165                 170                 175
Pro Lys Tyr Ala Leu Ala Leu Gln Glu Gln Leu Arg Ser Gln Ala Ser
            180                 185                 190
Asn Ser Ser Met Asn Asp Thr Ala Asp Ser Leu Asp Ser Cys Thr Ala
        195                 200                 205
Leu Gly Ala Leu Val Gly Ser Ser Lys Val Phe Phe Ser Cys Met Gln
    210                 215                 220
Ile Ser Met Thr Pro Met Ser Val Ser Met Pro Thr Val Met Pro Asn
225                 230                 235                 240
Thr Ser Gly Cys His
                245
```

(2) INFORMATION FOR SEQ ID NO:142:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 367 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...367

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

```
Met Ile Lys Ser Val Glu Ile Glu Asn Tyr Lys Asn Phe Glu His Leu
```

```
1               5                   10                  15
Lys Met Glu Asn Phe Lys Leu Ile Asn Phe Phe Thr Gly Gln Asn Asp
            20                  25                  30
Ala Gly Lys Thr Asn Leu Leu Glu Ala Leu Tyr Thr Asn Thr Gly Leu
        35                  40                  45
Cys Asp Pro Thr Ala Asn Gln Val Ser Leu Pro Pro Glu His Ala Val
    50                  55                  60
Asn Ile Ser Glu Phe Arg Lys Ile Lys Leu Asp Ala Asp Asn Leu Lys
65                  70                  75                  80
Thr Phe Phe Tyr Gln Gly Asn Thr Ala Asn Pro Ile Ser Ile Arg Thr
                85                  90                  95
Glu Phe Glu His Ala Thr Ile Pro Leu Thr Ile Gln Tyr Pro Thr Gln
            100                 105                 110
Thr Ser Tyr Ser Lys Asp Ile Asn Leu Asn Ser Asp Asp Ala His Met
        115                 120                 125
Thr Asn Leu Ile Asn Thr Thr Ile Thr Lys Pro Gln Leu Gln Phe Ser
    130                 135                 140
Tyr Asn Pro Ser Leu Ser Pro Met Thr Met Thr Tyr Glu Phe Glu Arg
145                 150                 155                 160
Gln Asn Leu Gly Leu Ile His Ser Asn Leu Asp Lys Ile Ala Gln Thr
                165                 170                 175
Tyr Lys Glu Asn Ala Met Phe Ile Pro Ile Glu Leu Ser Ile Val Asn
            180                 185                 190
Ser Leu Lys Ala Leu Glu Asn Leu Gln Leu Ala Ser Lys Glu Lys Glu
        195                 200                 205
Leu Ile Glu Ile Leu Gln Cys Phe Asn Pro Asn Ile Leu Asn Ala Asn
    210                 215                 220
Thr Ile Arg Lys Ser Val Tyr Ile Gln Ile Lys Asp Glu Asn Thr Pro
225                 230                 235                 240
Leu Glu Glu Ser Pro Lys Arg Leu Leu Asn Leu Phe Gly Trp Gly Phe
                245                 250                 255
Ile Lys Phe Phe Ile Met Val Ser Ile Leu Ile Asp Asn Arg Val Lys
            260                 265                 270
Tyr Leu Phe Ile Asp Glu Ile Glu Ser Gly Leu His His Thr Lys Met
        275                 280                 285
Gln Glu Phe Leu Lys Ala Leu Phe Lys Leu Ala Gln Lys Leu Gln Ile
    290                 295                 300
Gln Ile Phe Ala Thr Thr His Asn Lys Glu Phe Leu Leu Asn Ala Ile
305                 310                 315                 320
Asn Thr Ile Ser Asp Asn Glu Thr Gly Val Phe Lys Asp Ile Ala Leu
                325                 330                 335
Phe Glu Leu Glu Lys Glu Ser Ala Ser Gly Phe Ile Arg His Ser Tyr
            340                 345                 350
Ser Met Leu Glu Lys Ala Leu Tyr Arg Gly Met Glu Val Arg Gly
        355                 360                 365
```

(2) INFORMATION FOR SEQ ID NO:143:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 409 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...409

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

```
Met Ser Leu Ile Arg Val Asn Gly Glu Ala Phe Lys Leu Ser Leu Glu
1               5                  10                  15
Ser Leu Glu Glu Asp Pro Phe Glu Thr Lys Glu Thr Leu Glu Thr Leu
            20                  25                  30
Glu Thr Leu Ile Lys Gln Thr Ser Val Val Leu Leu Ala Ala Gly Glu
        35                  40                  45
Ser Lys Arg Phe Ser Arg Ala Ile Lys Lys Gln Trp Leu Arg Ser His
    50                  55                  60
His Thr Pro Leu Trp Leu Ser Val Tyr Glu Ser Phe Lys Glu Ala Leu
65                  70                  75                  80
Asp Phe Lys Glu Val Ile Leu Val Val Ser Glu Leu Asp Tyr Val Tyr
                85                  90                  95
Ile Gln Arg His Tyr Pro Lys Ile Lys Leu Val Lys Gly Gly Ala Ser
            100                 105                 110
Arg Gln Glu Ser Val Arg Asn Ala Leu Lys Val Ile Asp Ser Thr Tyr
        115                 120                 125
Thr Ile Thr Ser Asp Val Ala Arg Gly Leu Ala Asn Met Glu Ala Leu
    130                 135                 140
Lys Ser Leu Phe Leu Thr Leu Gln Gln Thr Ser His Tyr Cys Ile Ala
145                 150                 155                 160
Pro Tyr Leu Pro Cys Tyr Asp Thr Ala Ile Tyr Tyr Asn Glu Ala Leu
                165                 170                 175
Asp Arg Glu Ala Ile Lys Leu Ile Gln Thr Pro Gln Leu Ser His Thr
            180                 185                 190
Lys Thr Leu Gln Ser Ala Leu Asn Gln Gly Gly Phe Lys Asp Glu Ser
        195                 200                 205
Ser Ala Ile Leu Gln Ala Phe Pro Asn Ser Val Ser Tyr Ile Glu Gly
    210                 215                 220
Ser Lys Asp Leu His Lys Leu Thr Thr Ser Gly Asp Leu Lys Phe Phe
225                 230                 235                 240
Thr Pro Phe Phe Asn Pro Ala Lys Asp Thr Phe Ile Gly Met Gly Phe
                245                 250                 255
Asp Thr His Ala Phe Ile Lys Asp Lys Pro Met Val Leu Gly Gly Val
            260                 265                 270
Val Leu Asp Cys Glu Phe Gly Leu Lys Ala His Ser Asp Gly Asp Ala
        275                 280                 285
Leu Leu His Ala Val Ile Asp Ala Ile Leu Gly Ala Ile Lys Gly Gly
    290                 295                 300
Asp Ile Gly Glu Trp Phe Pro Asp Asn Asp Pro Lys Tyr Lys Asn Ala
305                 310                 315                 320
Ser Ser Lys Glu Leu Leu Lys Ile Val Leu Asp Phe Ser Gln Ser Ile
                325                 330                 335
Gly Phe Glu Leu Leu Glu Met Gly Ala Thr Ile Phe Ser Glu Ile Pro
            340                 345                 350
```

```
Lys Ile Thr Pro Tyr Lys Pro Ala Ile Leu Glu Asn Leu Ser Gln Leu
        355                 360                 365
Leu Gly Leu Glu Lys Ser Gln Ile Ser Leu Lys Ala Thr Thr Met Glu
    370                 375                 380
Lys Met Gly Phe Ile Gly Lys Gln Glu Gly Leu Leu Val Gln Ala His
385                 390                 395                 400
Val Ser Met Arg Tyr Lys Gln Lys Leu
                405
```

(2) INFORMATION FOR SEQ ID NO:144:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 270 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...270

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

```
Met Lys Lys Phe Val Ala Leu Gly Leu Leu Ser Ala Val Leu Ser Ser
1               5                   10                  15
Ser Leu Leu Ala Glu Gly Asp Gly Val Tyr Ile Gly Thr Asn Tyr Gln
            20                  25                  30
Leu Gly Gln Ala Arg Leu Asn Ser Asn Ile Tyr Asn Thr Gly Asp Cys
        35                  40                  45
Thr Gly Ser Val Val Gly Cys Pro Pro Gly Leu Thr Ala Asn Lys His
    50                  55                  60
Asn Pro Gly Gly Thr Asn Ile Asn Trp His Ser Lys Tyr Ala Asn Gly
65                  70                  75                  80
Ala Leu Asn Gly Phe Gly Leu Asn Val Gly Tyr Lys Lys Phe Phe Gln
                85                  90                  95
Phe Lys Ser Leu Asp Met Thr Ser Lys Trp Phe Gly Phe Arg Val Tyr
            100                 105                 110
Gly Leu Phe Asp Tyr Gly His Ala Asp Leu Gly Lys Gln Val Tyr Ala
        115                 120                 125
Pro Asn Lys Ile Gln Leu Asp Met Val Ser Trp Gly Val Gly Ser Asp
    130                 135                 140
Leu Leu Ala Asp Ile Ile Asp Lys Asp Asn Ala Ser Phe Gly Ile Phe
145                 150                 155                 160
Gly Gly Val Ala Ile Gly Gly Asn Thr Trp Lys Ser Ser Ala Ala Asn
                165                 170                 175
Tyr Trp Lys Glu Gln Ile Ile Glu Ala Lys Gly Pro Asp Val Cys Thr
            180                 185                 190
Pro Thr Tyr Cys Asn Pro Asn Ala Pro Tyr Ser Thr Asn Thr Ser Thr
        195                 200                 205
```

```
Val Ala Phe Gln Val Trp Leu Asn Phe Gly Val Arg Ala Asn Ile Tyr
    210                 215                 220
Lys His Asn Gly Val Glu Phe Gly Val Arg Val Pro Leu Leu Ile Asn
225                 230                 235                 240
Lys Phe Leu Ser Ala Gly Pro Asn Ala Thr Asn Leu Tyr Tyr His Leu
                245                 250                 255
Lys Arg Asp Tyr Ser Leu Tyr Leu Gly Tyr Asn Tyr Thr Phe
            260                 265                 270
```

(2) INFORMATION FOR SEQ ID NO:145:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 438 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...438

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

```
Met Ala Tyr Lys Pro Asn Lys Lys Lys Leu Lys Glu Leu Arg Glu Gln
1               5                   10                  15
Pro Asn Leu Phe Ser Ile Leu Asp Lys Gly Asp Val Ala Thr Asn Asn
            20                  25                  30
Pro Val Glu Glu Ser Asp Lys Ala Asn Lys Ile Gln Glu Pro Leu Pro
        35                  40                  45
Tyr Val Val Lys Thr Gln Ile Asn Lys Ala Ser Met Ile Ser Arg Asp
    50                  55                  60
Pro Ile Glu Trp Ala Lys Tyr Leu Ser Phe Glu Lys Arg Val Tyr Lys
65                  70                  75                  80
Asp Asn Ser Lys Glu Asp Val Asn Phe Phe Ala Asn Gly Glu Ile Lys
                85                  90                  95
Glu Ser Ser Arg Val Tyr Glu Ala Asn Lys Glu Gly Phe Glu Arg Arg
            100                 105                 110
Ile Thr Lys Arg Tyr Asp Leu Ile Asp Arg Asn Ile Asp Arg Asn Arg
        115                 120                 125
Glu Phe Phe Ile Lys Glu Ile Glu Ile Leu Thr His Thr Asn Ser Leu
    130                 135                 140
Lys Glu Leu Lys Glu Gln Gly Leu Glu Ile Gln Leu Thr His His Asn
145                 150                 155                 160
Glu Thr His Lys Lys Ala Leu Glu Asn Gly Asn Glu Ile Val Lys Glu
                165                 170                 175
Tyr Asp His Leu Lys Asp Ile Tyr Gln Glu Val Glu Arg Thr Lys Asp
            180                 185                 190
Gly Gly Leu Val Arg Glu Ile Ile Pro Ser Ile Ser Ser Ala Glu Tyr
        195                 200                 205
```

```
Phe Lys Leu Tyr Asn Lys Leu Pro Phe Glu Ser Ile Asn Asn Glu Asn
    210                 215                 220
Thr Lys Leu Asn Thr Asn Asp Asn Glu Glu Val Lys Lys Leu Glu Phe
225                 230                 235                 240
Glu Leu Ala Lys Glu Val His Ile Leu Ile Leu Glu Gln Gln Leu Leu
                245                 250                 255
Ser Ala Thr Asn Tyr Tyr Ser Trp Ile Asp Lys Asp Asp Asn Ala Asn
            260                 265                 270
Phe Ala Trp Lys Met His Arg Leu Ile Asn Glu Asn Lys Leu Lys Glu
        275                 280                 285
Asn His Leu Ser Ala Asn Asn Ala Asn Lys Ile Lys Gln Phe Phe Phe
    290                 295                 300
Asn Asn Gly Ser Ile Leu Gly Trp Thr Lys Glu Glu Gln Ser Ala Ile
305                 310                 315                 320
Gln Glu Asn Arg Asp Tyr Ser Leu Arg Ser Ala Leu Leu Ser Leu Glu
                325                 330                 335
Glu Ile Ala Gln Ala Lys Ile Glu Leu Gln Lys Tyr Tyr Glu Ser Val
            340                 345                 350
Tyr Val Asn Gly Asp Gly Asn Lys Arg Glu Ile Lys Pro Phe Lys Glu
        355                 360                 365
Ile Leu Arg Asp Thr Asn Asn Phe Glu Lys Ala Tyr Lys Glu Arg Tyr
    370                 375                 380
Asp Lys Leu Val Ser Leu Ser Ala Ala Ile Ile Gln Ala Lys Glu Gly
385                 390                 395                 400
Gly Asn Glu Arg Pro Asn Ser Ser Ala Asn Asn Asn Asn Pro Ile Lys
                405                 410                 415
Asn Thr Ile Glu Thr Asn Thr Ser Asn Asn Ile Ile Gln Asn Asn Asp
            420                 425                 430
Asn Ile Ile Ile Gln Ile
        435
```

(2) INFORMATION FOR SEQ ID NO:146:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 215 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...215

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

```
Met Gln Ala Leu Lys Ser Leu Leu Glu Val Ile Thr Lys Leu Gln Asn
1               5                   10                  15
Leu Gly Gly Tyr Leu Met His Ile Ala Ile Phe Ile Ile Phe Ile Trp
            20                  25                  30
```

Ile Gly Gly Leu Lys Phe Val Pro Tyr Glu Ala Glu Gly Ile Ala Pro
35 40 45
Phe Val Ala Asn Ser Pro Phe Phe Ser Phe Met Tyr Lys Phe Glu Lys
50 55 60
Pro Ala Tyr Lys Gln His Lys Met Ser Glu Ser Gln Ser Met Gln Glu
65 70 75 80
Glu Met Gln Asp Asn Pro Lys Ile Val Glu Asn Lys Glu Trp His Lys
85 90 95
Glu Asn Arg Thr Tyr Leu Val Ala Glu Gly Leu Gly Ile Thr Ile Met
100 105 110
Ile Leu Gly Ile Leu Val Leu Leu Gly Leu Trp Met Pro Leu Met Gly
115 120 125
Val Val Gly Gly Leu Leu Val Ala Gly Met Thr Ile Thr Thr Leu Ser
130 135 140
Phe Leu Phe Thr Thr Pro Glu Val Phe Val Asn Gln His Phe Pro Trp
145 150 155 160
Leu Ser Gly Ala Gly Arg Leu Val Val Lys Asp Leu Ala Leu Phe Ala
165 170 175
Gly Gly Leu Phe Val Ala Gly Phe Asp Ala Lys Arg Tyr Leu Glu Gly
180 185 190
Lys Gly Phe Cys Leu Met Asp Arg Ser Ser Val Gly Ile Lys Thr Lys
195 200 205
Cys Ser Ser Gly Cys Cys Ser
210 215

(2) INFORMATION FOR SEQ ID NO:147:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

TATACCATGG TGGGCGCTAA 20

(2) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

ATGAATTCGA GTAAGGATTT TTG 23

(2) INFORMATION FOR SEQ ID NO:149:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

TTAACCATGG TGAAAAGCGA TA 22

(2) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

TAGAATTCGC ATAACGATCA ATC 23

(2) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

ATATCCATGG TGAGTTTGAT GA 22

(2) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

ATGAATTCAA TTTTTTATTT TGCCA 25

(2) INFORMATION FOR SEQ ID NO:153:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

AATTCCATGG TGGGGGCTAT G 21

(2) INFORMATION FOR SEQ ID NO:154:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

ATGAATTCTC GATAGCCAAA ATC 23

(2) INFORMATION FOR SEQ ID NO:155:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

AATTCCATGG TGCATAACTT CCATT 25

(2) INFORMATION FOR SEQ ID NO:156:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

AAGAATTCTC TAGCATCCAA ATGGA 25

(2) INFORMATION FOR SEQ ID NO:157:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

ATTTCCATGG TCATGTCTCA TATT 24

(2) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

ATGAATTCCA TCTTTTATTC CAC 23

(2) INFORMATION FOR SEQ ID NO:159:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

AACCATGGTG ATTTTAAGCA TTGAAAG 27

(2) INFORMATION FOR SEQ ID NO:160:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

AAGAATTCCA CTCAAAATTT TTTAACAG 28

(2) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

GATCATCCAT ATGTTATCTT CTAAT 25

(2) INFORMATION FOR SEQ ID NO:162:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

TGAATTCAAC CATTTTAACC CTG 23

(2) INFORMATION FOR SEQ ID NO:163:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

TATACCATGG TGAAATTTTT TCTTTTA 27

(2) INFORMATION FOR SEQ ID NO:164:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

AGAATTCAAT TGCGTCTTGT AAAAG 25

(2) INFORMATION FOR SEQ ID NO:165:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

TATACCATGG TGATGGACAA ACTC 24

(2) INFORMATION FOR SEQ ID NO:166:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

ATGAATTCCC ACTTGGGGCG ATA 23

(2) INFORMATION FOR SEQ ID NO:167:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

TTATGGATCC AAACCAATTA AAACT 25

(2) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

TATCTCGAGT TATAGAGAAG GGC 23

(2) INFORMATION FOR SEQ ID NO:169:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

TTAACCATGG TGAAAAGCGA TA 22

(2) INFORMATION FOR SEQ ID NO:170:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

TAGAATTCGC CTCTAAAACT TTAG 24

(2) INFORMATION FOR SEQ ID NO:171:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

TTAACCATGG TGAAAAGCGA TA 22

(2) INFORMATION FOR SEQ ID NO:172:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

TAGAATTCGC ATAACGATCA ATC 23

(2) INFORMATION FOR SEQ ID NO:173:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

ATATCCATGG TGAGTTTGAT GA 22

(2) INFORMATION FOR SEQ ID NO:174:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

ATGAATTCAA TTTTTTATTT TGCCA 25

(2) INFORMATION FOR SEQ ID NO:175:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

AATTCCATGG CTATCCAAAT CCG 23

(2) INFORMATION FOR SEQ ID NO:176:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

ATGAATTCGC CAAAATCGTA GTATT 25

(2) INFORMATION FOR SEQ ID NO:177:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:177:

GATACCATGG AATTTATGAA AAAG 24

(2) INFORMATION FOR SEQ ID NO:178:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

TGAATTCGAA AAAGTGTAGT TATAC 25

(2) INFORMATION FOR SEQ ID NO:179:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

CCCTTCATTT TAGAAATCG 19

(2) INFORMATION FOR SEQ ID NO:180:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

ATTTCAACCA ATTCAATGCG 20

(2) INFORMATION FOR SEQ ID NO:181:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

GCCCCTTTTG ATTTGAAGCT 20

(2) INFORMATION FOR SEQ ID NO:182:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

TCGCTCCAAG ATACCAAGAA GT 22

(2) INFORMATION FOR SEQ ID NO:183:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

CTTGAATTAG GGGCAAAGAT CG 22

(2) INFORMATION FOR SEQ ID NO:184:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

ATGCGTTTTT ACCCAAAGAA GT 22

(2) INFORMATION FOR SEQ ID NO:185:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

ATAACGCCAC TTCCTTATTG GT 22

(2) INFORMATION FOR SEQ ID NO:186:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

CTTTGGGTAA AAACGCATC 19

(2) INFORMATION FOR SEQ ID NO:187:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

CGATCTTTGA TCCTAATTCA 20

(2) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

ATCAAGTTGC CTATGCTGA 19

(2) INFORMATION FOR SEQ ID NO:189:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

TTGAACACTT TTGATTATGC GG 22

(2) INFORMATION FOR SEQ ID NO:190:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

GGATTATGCG ATTGTTTTAC AAG 23

(2) INFORMATION FOR SEQ ID NO:191:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

GTCTTTAGCA AAAATGGCGT C 21

(2) INFORMATION FOR SEQ ID NO:192:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

AATGAGCGTA AGAGAGCCTT C 21

(2) INFORMATION FOR SEQ ID NO:193:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

CTTATGGGGG TATTGTCA 18

(2) INFORMATION FOR SEQ ID NO:194:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

AGCATGTGGG TATCCAGC 18

(2) INFORMATION FOR SEQ ID NO:195:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

AGGTTGTTGC CTAAAGACT 19

(2) INFORMATION FOR SEQ ID NO:196:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

CTGCCTCCAC CTTTGATC 18

(2) INFORMATION FOR SEQ ID NO:197:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:

ACCAATATCA ATTGGCACT 19

(2) INFORMATION FOR SEQ ID NO:198:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

ACTTGGAAAA GCTCTGCA 18

(2) INFORMATION FOR SEQ ID NO:199:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:

CTTGCTTGTC ATATCTAGC 19

(2) INFORMATION FOR SEQ ID NO:200:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:200:

GTTGAAGTGT TGGTGCTA 18

(2) INFORMATION FOR SEQ ID NO:201:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:

CAAGCAAGTG GTTTGGTTTT AG 22

(2) INFORMATION FOR SEQ ID NO:202:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

TGGAAAGAGC AAATCATTGA AG 22

(2) INFORMATION FOR SEQ ID NO:203:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:

GCCCATAATC AAAAAGCCCA T 21

(2) INFORMATION FOR SEQ ID NO:204:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

CTAAAACCAA ACCACTTGCT TGTC 24

(2) INFORMATION FOR SEQ ID NO:205:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:205:

GTAAAACGAC GGCCAG 16

(2) INFORMATION FOR SEQ ID NO:206:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...17

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:206:

CAGGAAACAG CTATGAC 17

(2) INFORMATION FOR SEQ ID NO:207:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:207:

ATCTTACCTA TCACCTCAAA T 21

(2) INFORMATION FOR SEQ ID NO:208:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:208:

AGACAGCAAC ATCTTTGTGA A 21

## CLAIMS

1. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* polypeptide at least about 60% homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 74-SEQ ID NO: 146.

2. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* polypeptide selected from the group consisting of SEQ ID NO: 74-SEQ ID NO: 146.

3. An isolated nucleic acid which encodes an *H. pylori* polypeptide, comprising a nucleotide sequence at least about 60% homologous to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 73, or a complement thereof.

4. The isolated nucleic acid of claim 1, comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 73, or a complement thereof.

5. An isolated nucleic acid molecule encoding an *H. pylori* polypeptide, comprising a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 73, or a complement thereof.

6. An isolated nucleic acid comprising a nucleotide sequence of at least 8 nucleotides in length, wherein the sequence hybridizes under stringent hybridization conditions to a nucleic acid having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 73, or a complement thereof.

7. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cell envelope polypeptide or a fragment thereof, said nucleic acid selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 25, SEQ ID NO: 48, SEQ ID NO: 16, SEQ ID NO: 10, SEQ ID NO: 45, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 7, SEQ ID NO: 39, SEQ ID NO: 55, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 1, SEQ ID NO: 42, SEQ ID NO: 14, SEQ ID NO: 43, SEQ ID

NO: 11, SEQ ID NO: 71, SEQ ID NO: 17, SEQ ID NO: 57, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 8, and SEQ ID NO: 21, or a complement thereof.

8. The isolated nucleic acid of claim 7, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 25, and SEQ ID NO: 48, or a complement thereof.

9. The isolated nucleic acid of claim 7, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 16, SEQ ID NO: 10, SEQ ID NO: 45, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 7, SEQ ID NO: 39, SEQ ID NO: 55, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 1, SEQ ID NO: 42, SEQ ID NO: 14, SEQ ID NO: 43, SEQ ID NO: 11, and SEQ ID NO: 71, or a complement thereof.

10. The isolated nucleic acid of claim 9, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 42, SEQ ID NO: 14, SEQ ID NO: 43, SEQ ID NO: 11 and SEQ ID NO:71, or a complement thereof.

11. The isolated nucleic acid of claim 9, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 16, SEQ ID NO: 45, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 7, SEQ ID NO: 39, SEQ ID NO: 55, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, or a complement thereof.

12. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cell envelope polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 76, SEQ ID NO: 98, SEQ ID NO: 121, SEQ ID NO: 89, SEQ ID NO: 83, SEQ ID NO: 118, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 80, SEQ ID NO: 112, SEQ ID NO: 128, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO:

101, SEQ ID NO: 103, SEQ ID NO: 125, SEQ ID NO: 127, SEQ ID NO: 129, SEQ ID NO: 131, SEQ ID NO: 74, SEQ ID NO: 115, SEQ ID NO: 87, SEQ ID NO: 116, SEQ ID NO: 84, SEQ ID NO: 144, SEQ ID NO: 90, SEQ ID NO: 130, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 81, and SEQ ID NO: 94.

13. The isolated nucleic acid of claim 12, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 76, SEQ ID NO: 98, and SEQ ID NO: 121.

14. The isolated nucleic acid of claim 12, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 89, SEQ ID NO: 83, SEQ ID NO: 118, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 80, SEQ ID NO: 112, SEQ ID NO: 128, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 125, SEQ ID NO: 127, SEQ ID NO: 129, SEQ ID NO: 131, SEQ ID NO: 74, SEQ ID NO: 115, SEQ ID NO: 87, SEQ ID NO: 116, SEQ ID NO: 84, SEQ ID NO: 144, SEQ ID NO: 90, and SEQ ID NO: 130.

15. The isolated nucleic acid of claim 14, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof selected from the group consisting of SEQ ID NO: 74, SEQ ID NO: 115, SEQ ID NO: 87, SEQ ID NO: 116, and SEQ ID NO: 84 and SEQ ID NO:144.

16. The isolated nucleic acid of claim 14, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof selected from the group consisting of SEQ ID NO: 89, SEQ ID NO: 118, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 80, SEQ ID NO: 112, SEQ ID NO: 128, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 125, SEQ ID NO: 127, SEQ ID NO: 129, and SEQ ID NO: 131.

17. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* secreted polypeptide or a fragment thereof, said nucleic acid selected from the group consisting of SEQ ID NO: 72, SEQ ID NO: 32, SEQ ID NO: 51, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 22, SEQ ID NO: 29, SEQ

ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 49, SEQ ID NO: 53, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67, and SEQ ID NO: 68, or a complement thereof.

18. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* secreted polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 145, SEQ ID NO: 105, SEQ ID NO: 124, SEQ ID NO: 75, SEQ ID NO: 77, SEQ ID NO: 82, SEQ ID NO: 86, SEQ ID NO: 95, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 107, SEQ ID NO: 109, SEQ ID NO: 111, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 117, SEQ ID NO: 119, SEQ ID NO: 122, SEQ ID NO: 126, SEQ ID NO: 132, SEQ ID NO: 134, SEQ ID NO: 135, SEQ ID NO: 136, SEQ ID NO: 138, SEQ ID NO: 139, SEQ ID NO: 140, and SEQ ID NO: 141.

19. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cellular polypeptide or a fragment thereof, said nucleic acid selected from the group consisting of SEQ ID NO: 12, SEQ ID NO: 15, SEQ ID NO: 20, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 47, SEQ ID NO: 50, SEQ ID NO: 60, SEQ ID NO: 64, SEQ ID NO: 69, SEQ ID NO: 70, and SEQ ID NO: 73, or a complement thereof.

20. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cellular polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 85, SEQ ID NO: 88, SEQ ID NO: 93, SEQ ID NO: 96, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 100, SEQ ID NO: 120, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 137, SEQ ID NO: 142, SEQ ID NO: 143, and SEQ ID NO: 146.

21. A probe comprising a nucleotide sequence consisting of at least 8 nucleotides of a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 73, or a complement thereof.

22. A recombinant expression vector comprising the nucleic acid of any of claims 1, 2, 3, 4, 5, 6, 7, 12, 17, 18, 19 or 20 operably linked to a transcription regulatory element.

23. A cell comprising a recombinant expression vector of claim 22.

24. A method for producing an *H. pylori* polypeptide comprising culturing a cell of claim 23 under conditions that permit expression of the polypeptide.

25. The method of claim 24, further comprising purifying the polypeptide from the cell.

26. A method for detecting the presence of a *Helicobacter* nucleic acid in a sample comprising:

(a) contacting a sample with a nucleic acid of any of claims 6 or 21 so that a hybrid can form between the probe and a *Helicobacter* nucleic acid in the sample; and

(b) detecting the hybrid formed in step (a), wherein detection of a hybrid indicates the presence of a *Helicobacter* nucleic acid in the sample.

27. An isolated *H. pylori* polypeptide comprising an amino acid sequence at least about 60% homologous to an *H. pylori* polypeptide selected from the group consisting of SEQ ID NO: 74-SEQ ID NO: 146.

28. An isolated *H. pylori* polypeptide which is encoded by a nucleic acid comprising a nucleotide sequence at least about 60% homologous to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 73.

29. The isolated *H. pylori* polypeptide of claim 28, wherein said polypeptide is encoded by a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 73.

30. An isolated *H. pylori* polypeptide which is encoded by a nucleic acid which hybridizes under stringent hybridization conditions to a nucleic acid selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 73, or a complement thereof.

31. An isolated *H. pylori* polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 74-SEQ ID NO: 146.

32. An isolated *H. pylori* cell envelope polypeptide or a fragment thereof, wherein said polypeptide is selected from the group consisting of SEQ ID NO: 76, SEQ

ID NO: 98, SEQ ID NO: 121, SEQ ID NO: 89, SEQ ID NO: 83, SEQ ID NO: 118, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 80, SEQ ID NO: 112, SEQ ID NO: 128, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 125, SEQ ID NO: 127, SEQ ID NO: 129, SEQ ID NO: 131, SEQ ID NO: 74, SEQ ID NO: 115, SEQ ID NO: 87, SEQ ID NO: 116, SEQ ID NO: 84, SEQ ID NO: 144, SEQ ID NO: 90, SEQ ID NO: 130, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 81, and SEQ ID NO: 94.

33. The isolated polypeptide of claim 32, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 76, SEQ ID NO: 98, and SEQ ID NO: 121.

34. The isolated polypeptide of claim 32, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 89, SEQ ID NO: 83, SEQ ID NO: 118, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 80, SEQ ID NO: 112, SEQ ID NO: 128, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 125, SEQ ID NO: 127, SEQ ID NO: 129, SEQ ID NO: 131, SEQ ID NO: 74, SEQ ID NO: 115, SEQ ID NO: 87, SEQ ID NO: 116, SEQ ID NO: 84, SEQ ID NO: 144, SEQ ID NO: 90, and SEQ ID NO: 130.

35. The isolated polypeptide of claim 34, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof selected from the group consisting of SEQ ID NO: 74, SEQ ID NO: 115, SEQ ID NO: 87, SEQ ID NO: 116, and SEQ ID NO: 84 and SEQ ID NO:144.

36. The isolated polypeptide of claim 34, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof selected from the group consisting of SEQ ID NO: 89, SEQ ID NO: 118, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 80, SEQ ID NO: 112, SEQ ID NO: 128, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 125, SEQ ID NO: 127, SEQ ID NO: 129, and SEQ ID NO: 131.

37. An isolated *H. pylori* cell envelope polypeptide or a fragment thereof, wherein said polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 25, SEQ ID NO: 48, SEQ ID NO: 16, SEQ ID NO: 10, SEQ ID NO: 45, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 7, SEQ ID NO: 39, SEQ ID NO: 55, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 1, SEQ ID NO: 42, SEQ ID NO: 14, SEQ ID NO: 43, SEQ ID NO: 11, SEQ ID NO: 71, SEQ ID NO: 17, SEQ ID NO: 57, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 8, and SEQ ID NO: 21.

38. The isolated polypeptide of claim 37, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 25, and SEQ ID NO: 48.

39. The isolated polypeptide of claim 37, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 16, SEQ ID NO: 10, SEQ ID NO: 45, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 7, SEQ ID NO: 39, SEQ ID NO: 55, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 1, SEQ ID NO: 42, SEQ ID NO: 14, SEQ ID NO: 43, SEQ ID NO: 11, and SEQ ID NO: 71.

40. The isolated polypeptide of claim 39, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 42, SEQ ID NO: 14, SEQ ID NO: 43, SEQ ID NO: 11 and SEQ ID NO:71.

41. The isolated polypeptide of claim 39, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 16, SEQ ID NO: 45, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 7, SEQ ID NO: 39, SEQ ID NO: 55, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58.

42. An isolated *H. pylori* cellular polypeptide or a fragment thereof, wherein said polypeptide is selected from the group consisting of SEQ ID NO: 85, SEQ ID NO: 88, SEQ ID NO: 93, SEQ ID NO: 96, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 100, SEQ ID NO: 120, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 137, SEQ ID NO: 142, SEQ ID NO: 143, and SEQ ID NO: 146.

43. An isolated *H. pylori* cellular polypeptide or a fragment thereof, wherein said polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 12, SEQ ID NO: 15, SEQ ID NO: 20, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 47, SEQ ID NO: 50, SEQ ID NO: 60, SEQ ID NO: 64, SEQ ID NO: 69, SEQ ID NO: 70, and SEQ ID NO: 73.

44. An isolated *H. pylori* secreted polypeptide or a fragment thereof, wherein said polypeptide is selected from the group consisting of SEQ ID NO: 145, SEQ ID NO: 105, SEQ ID NO: 124, SEQ ID NO: 75, SEQ ID NO: 77, SEQ ID NO: 82, SEQ ID NO: 86, SEQ ID NO: 95, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 107, SEQ ID NO: 109, SEQ ID NO: 111, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 117, SEQ ID NO: 119, SEQ ID NO: 122, SEQ ID NO: 126, SEQ ID NO: 132, SEQ ID NO: 134, SEQ ID NO: 135, SEQ ID NO: 136, SEQ ID NO: 138, SEQ ID NO: 139, SEQ ID NO: 140, and SEQ ID NO: 141.

45. An isolated *H. pylori* secreted polypeptide or a fragment thereof, wherein said polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 72, SEQ ID NO: 32, SEQ ID NO: 51, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 22, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 49, SEQ ID NO: 53, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67, and SEQ ID NO: 68.

46. A fusion protein comprising an *H. pylori* polypeptide which comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 74-SEQ ID NO: 146 operatively linked to a non-*H. pylori* polypeptide.

47. A vaccine formulation for prophylaxis or treatment of an *H. pylori* infection comprising an effective amount of at least one isolated nucleic acid of any of claims 1, 2, 3, 4, 5, 6, 7, 12, 17, 18, 19, or 20.

48. A vaccine formulation for prophylaxis or treatment of an *H. pylori* infection comprising an effective amount of at least one *H. pylori* polypeptide or a fragment thereof of any of claims 26, 27, 28, 29, 30, 31, 32, 37, 42, 43, 44 or 45.

49. A vaccine formulation of claim 47, further comprising a pharmaceutically acceptable carrier.

50. A vaccine formulation of claim 48, further comprising a pharmaceutically acceptable carrier.

51. A vaccine formulation of claim 49, wherein the pharmaceutically acceptable carrier comprises an adjuvant.

52. A vaccine formulation of claim 50, wherein the pharmaceutically acceptable carrier comprises an adjuvant.

53. A vaccine formulation of claim 49, wherein the pharmaceutically acceptable carrier comprises a delivery system.

54. A vaccine formulation of claim 50, wherein the pharmaceutically acceptable carrier comprises a delivery system.

55. A vaccine formulation of claim 53, wherein the delivery system comprises a live vector.

56. A vaccine formulation of claim 54, wherein the delivery system comprises a live vector.

57. A vaccine formulation of claim 55, wherein the live vector is a bacteria or a virus.

58. A vaccine formulation of claim 56, wherein the live vector is a bacteria or a virus.

59. A vaccine formulation of claim 53, wherein the pharmaceutically acceptable carrier further comprises an adjuvant.

60. A vaccine formulation of claim 54, wherein the pharmaceutically acceptable carrier further comprises an adjuvant.

61. A method of treating or reducing a risk of *H. pylori* infection in a subject comprising administering to a subject a vaccine formulation of claim 47, such that treatment or reduction of risk of *H. pylori* infection occurs.

62. A method of treating or reducing a risk of *H. pylori* infection in a subject comprising administering to a subject a vaccine formulation of claim 48, such that treatment or reduction of risk of *H. pylori* infection occurs.

63. A method of producing a vaccine formulation comprising: combining at least one isolated *H. pylori* polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 74-SEQ ID NO: 146 with a pharmaceutically acceptable carrier to thereby form a vaccine formulation.

64. A method of producing a vaccine formulation comprising:

(a) providing at least one isolated *H. pylori* polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 74-SEQ ID NO: 146; and

(b) combining at least one said isolated *H. pylori* polypeptide or a fragment thereof with a pharmaceutically acceptable carrier to thereby form a vaccine formulation.

65. A method of producing a vaccine formulation comprising:

(a) culturing a cell under condition that permit expression of an *H. pylori* polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 74-SEQ ID NO: 146;

(b) isolating said *H. pylori* polypetide from said cell; and

(c) combining at least one said isolated *H. pylori* polypeptide or a fragment thereof with a pharmaceutically acceptable carrier to thereby form a vaccine formulation.

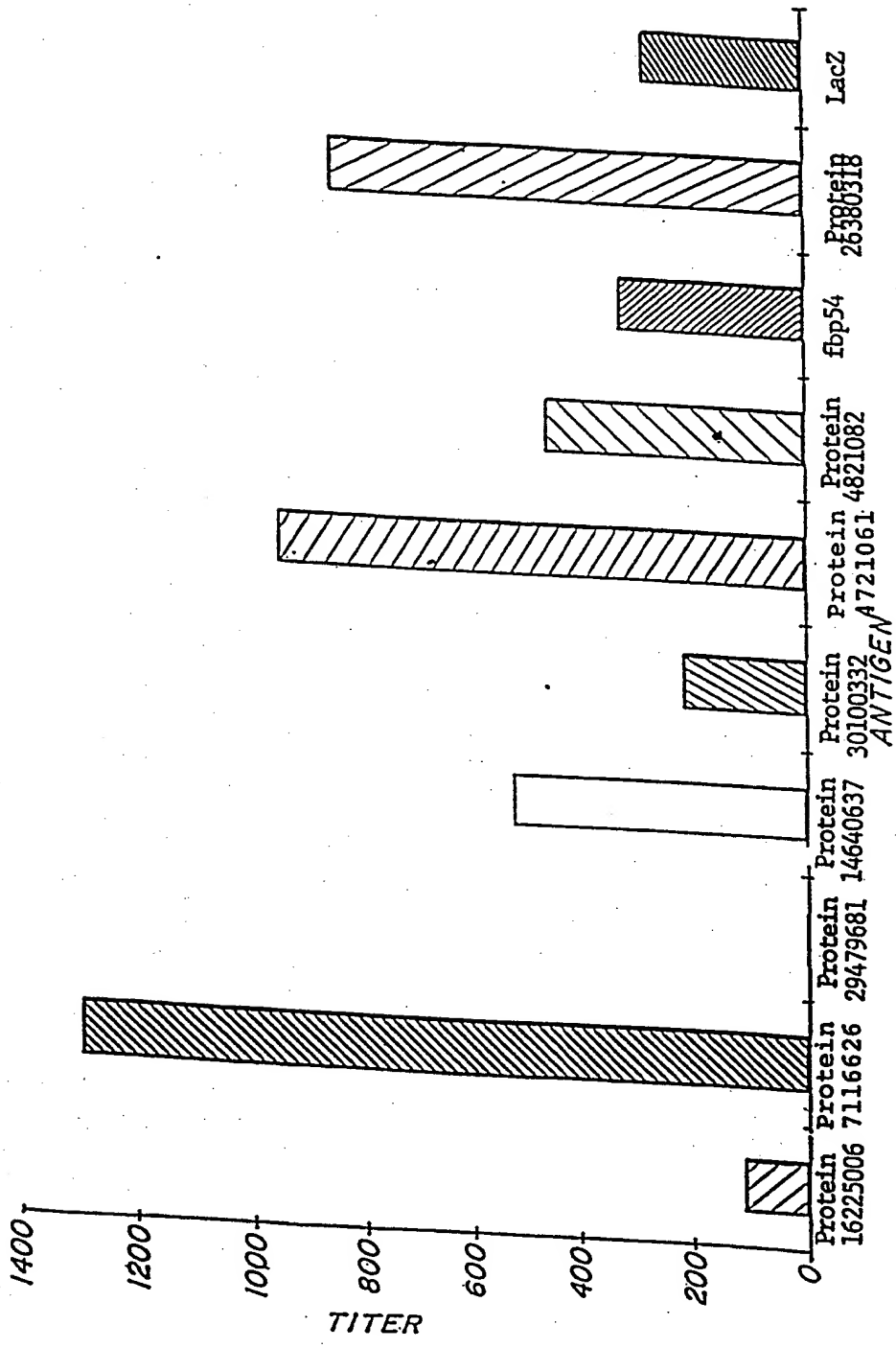
1400
1200
1000
800
600
400
200
0
TITER
Protein 16225006
Protein 7116626
Protein 29479681
Protein 14640637
Protein 30100332
Protein 4721061
Protein 4821082
fbp54
Protein 26380318
LacZ
ANTIGEN

FIG. 1

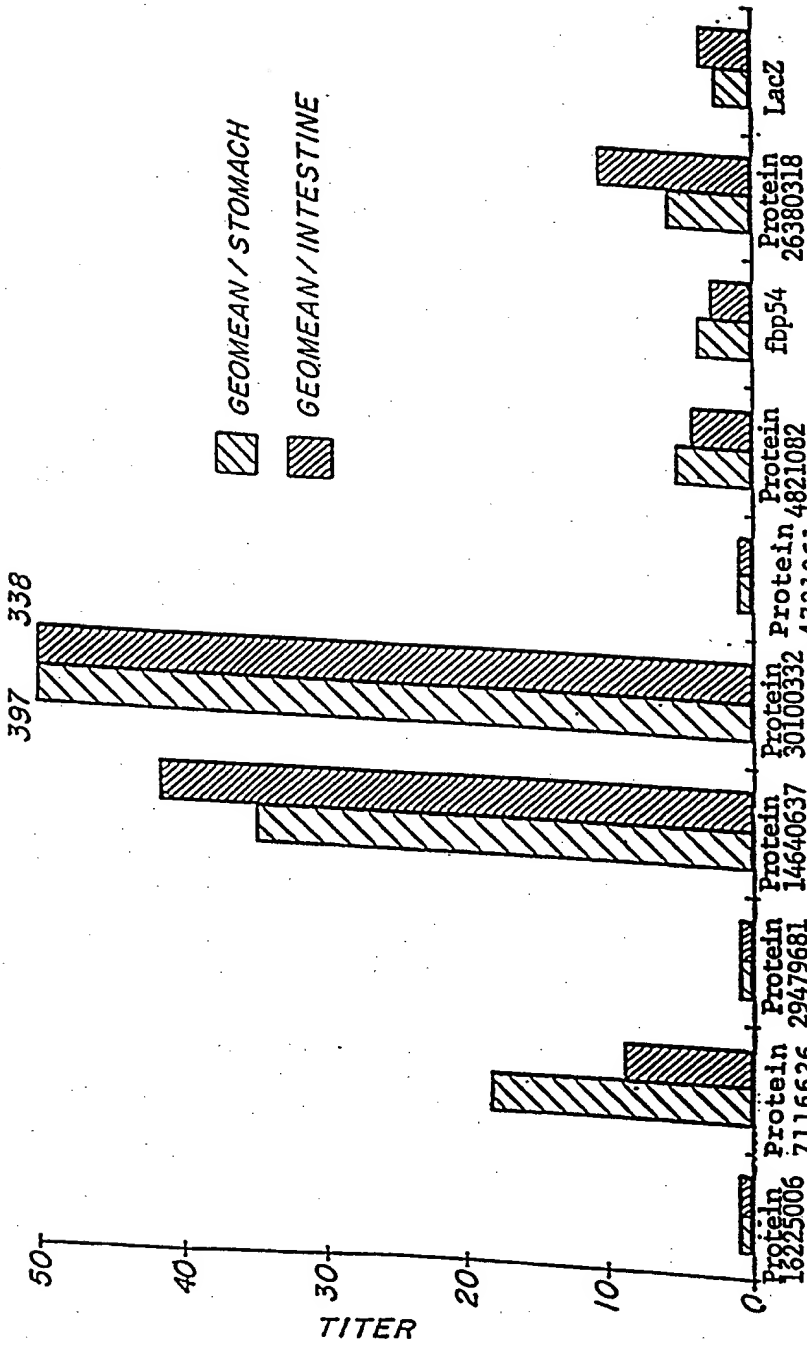
GEOMEAN / STOMACH
GEOMEAN / INTESTINE
397
338
TITER
50
40
30
20
10
0
Protein 13225006
Protein 7116626
Protein 2947968
Protein 14640637
Protein 30100332
Protein 4721061
Protein 4821082
fbp54
Protein 26380318
LacZ
ANTIGEN

FIG. 2

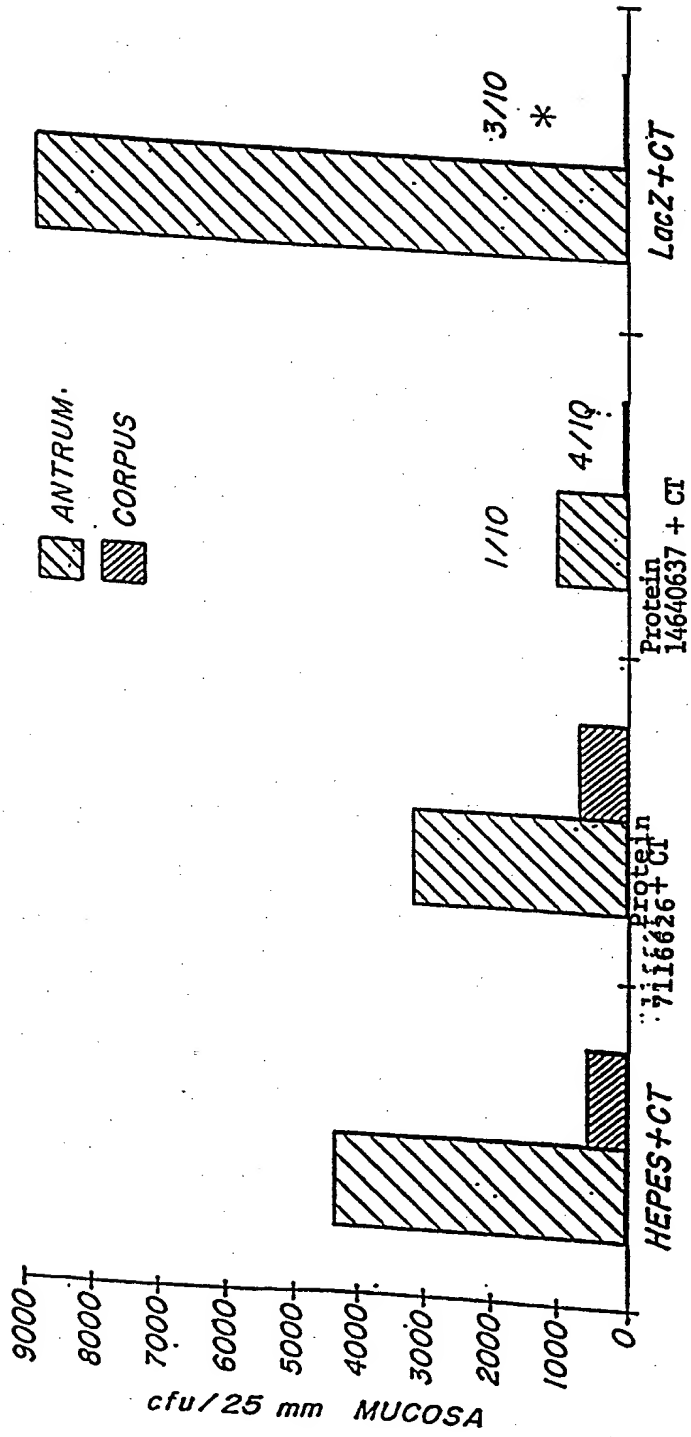
ANTRUM.
CORPUS
3/10
*
1/10
4/10
HEPES+CT
Protein 7116626 + CT
Protein 14640637 + CT
LacZ+CT
TREATMENT
cfu/25 mm MUCOSA
0
1000
2000
3000
4000
5000
6000
7000
8000
9000

FIG. 3

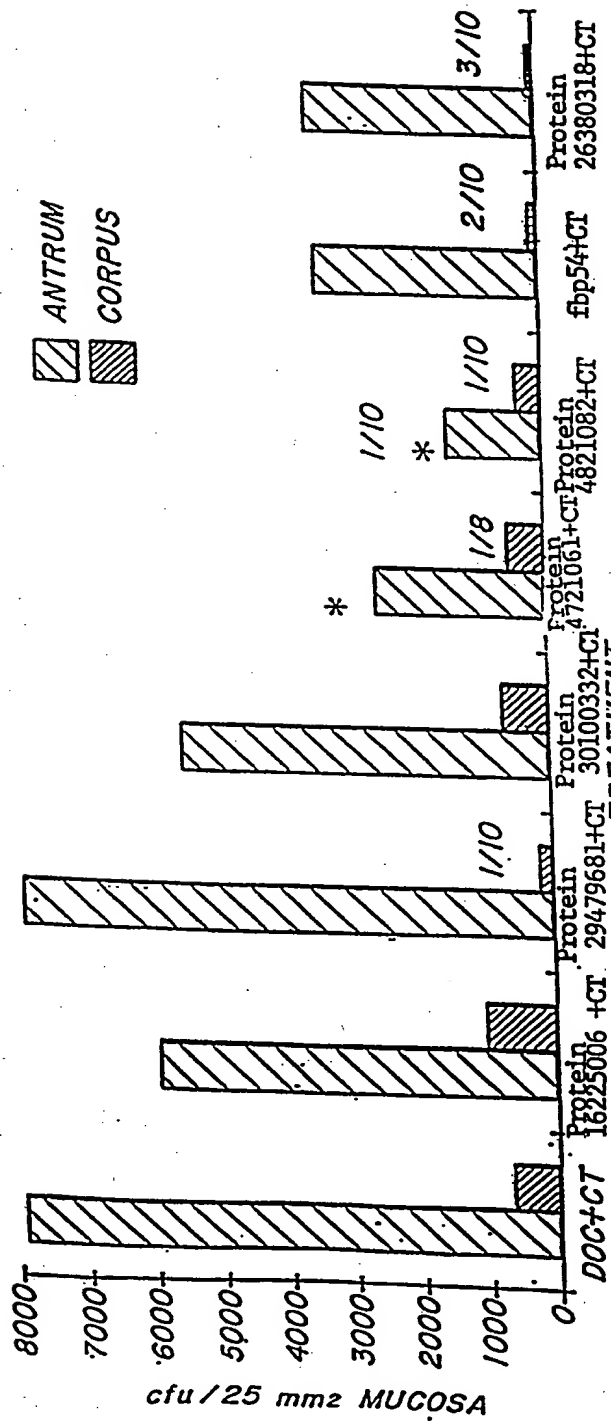
ANTRUM
CORPUS
8000
7000
6000
5000
4000
3000
2000
1000
0
cfu / 25 mm2 MUCOSA
DOC+CT
Protein 16225006 +CT
Protein 29479681+CT
Protein 30100332+CT
Protein 47210611+CT
Protein 48210824+CT
fbp54+CT
Protein 26380318+CT
1/10
1/10
1/8
1/10
2/10
3/10
*
*
TREATMENT

FIG. 4

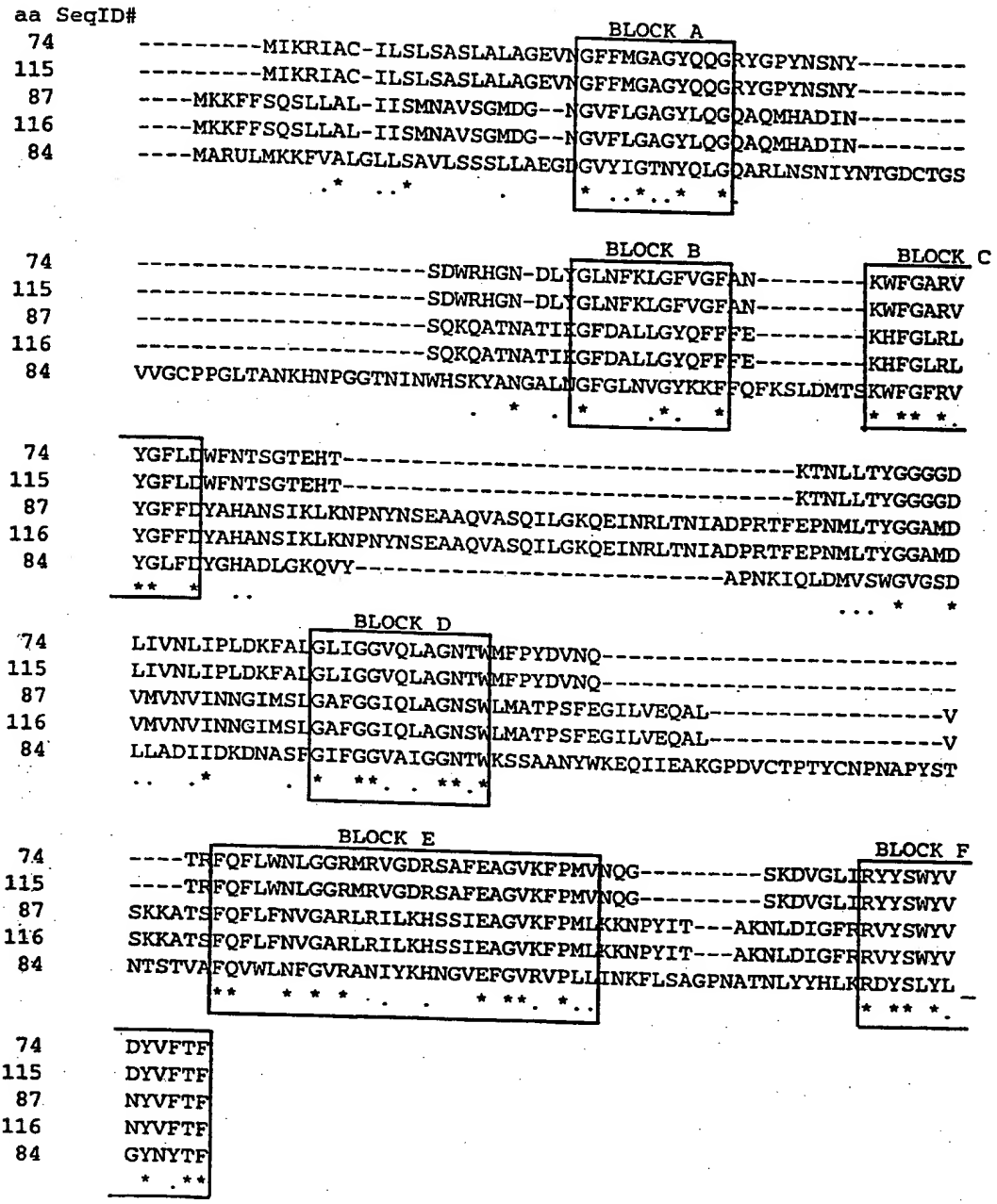


FIGURE 5

```
aaSeqID#
 83      MRKLFIPLLLFSALEANEKNGFFIEAGFETGLLEGTQTQEKRHTTTKNTYATYNYLPTDT
 89      ------------------------------------------------------------
108      MRKLFIPLLLFSALEANEKNGFFIEAGFETGLLEGTQTQEKRHTTTKNTYATYNYLPTDT
118      MRKLFIPLLLFSALEANEKNGFFIEAGFETGLLEGTQTQEKRHTTTKNTYATYNYLPTDT
         ************************************************************

 83      ILKRAANLFTNAEAISKLKFSSLSPVRVLYMYNGQLTIENFLPYNLNNVKLSFTDAQGNV
 89      ------------------------------------------------------------
108      ILKRAANLFTNAEAISKLKFSSLSPVRVLYMYNGQLTIENFLPYNLNNVKLSFTDAQGNT
118      ILKRAANLFTNAEAISKLKFSSLSPVRVLYMYNGQLTIENFLPYNLNNVKLSFTDAQGNV
         ************************************************************

 83      IDLGVIETIPKHSKIVLPGEAFDSL-----KIDPYTLFLPKIEATSTSISDANTQRVFET
 89      ----VIETIPKHSKIVLPGEAFDSLKEAFDKIDPYTFFFPKFEATSTSISDTNTQRVFET
108      IDLGVIETIPKHSKIVLPGEAFDSLKEAFDKIDPYTLFLPKFEATSTSISDTNTQRVFET
118      IDLGVIETIPKHSKIVLPGEAFDSL-----KIDPYTLFLPKIEATSTSISDANTQRVFET
         *************************     *********** ********* ********

 83      LNKIKTNLVVNYRNEN-----------KFKDHENHWEAFTPQTAEEFTNLMLNMIAVLDS
 89      LNNIKTNLIMKYSNENPNNFNTCPYNNNGNTKNDCWQNFTPQTAEEFTNLMLNMIAVLDS
108      LNNIKTNLIMKYSNENPNNFNTCPYNNNGNTKNDCWQNFTPQTAEEFTNLMLNMIAVLDS
118      LNKIKTNLVVNYRNEN-----------KFKDHENHWEAFTPQTAEEFTNLMLNMIAVLDS
         ** *****.. * ***              .   *. ************************

 83      QSWGDAILNAPFEFTNSPTDCDNDPSKCVNPGTNGLVNSKVDQKYVLNKQDIVNKFKNKA
 89      QSWGDAILNAPFEFTNSSTDCDSDPSKCVNPGVNGRVDTKVDQQYILNKQGIINNFRKKI
108      QSWGDAILNAPFEFTNSSTDCDSDPSKCVNPGVNGRVDTKVDQQYILNKQGIINNFRKKI
118      QSWGDAILNAPFEFTNSPTDCDNDPSKCVNPGTNGLVNSKVDQKYVLNKQDIVNKFKNKA
         ****************** **** ********* ** * .****.*.**** *.* *. *

 83      DLDVIVLKDSGVVGLGSDITPSNNDDGKHYGQLGVVASALDPKKLFGDNLKTINLEDLRT
 89      EIDAVVLKNSGVVGLANGYG----NDG-EYGTLGVEAYALDPKKLFGNDLKTINLEDLRT
108      EIDAVVLKNSGVVGLANGYG----NDG-EYGTLGVEAYALDPKKLFGNDLKTINLEDLRT
118      DLDVIVLKDSGVVGLGSDITPSNNDDGKHYGQLGVVASALDPKKLFGDNLKTINLEDLRT
         ..* .*** ******          **  ** *** * *********  ***********

 83      ILHEFSHTKGYGHNGNMTYQRVPVTKDGQVEKDSNGKPKDSDGLPYNVC-----------
 89      ILHEFSHTKGYGHNGNMTYQRVPVTKDGQVEKDSNGKPKDSDGLPYNVCSLYGGSNQPAF
108      ILHEFSHTKGYGHNGNMTYQRVPVTKDGQVEKDSNGKPKDSDGLPYNVCSLYGGSNQPAF
118      ILHEFSHTKGYGHNGNMTYQRVPVTKDGQVEKDSNGKPKDSDGLPYNVCSLYGGSNQPAF
         ************************************************************

 83      ------------------------------------------------------------
 89      PSNYPNSIYHNCADVPAGFLGVTAAVWQQLINQNALPINYANLGSQTNYNLNASLNTQDL
108      PSNYPNSIYHNCADVPAGFLGVTAAVWQQLINQNALPINYANLGSQTNYNLNASLNTQDL
118      PSNYPNSIYHNCADVPAGFLGVTAAVWQQLINQNALPINYANLGSQTNYNLNASLNTQDL
         ************************************************************
```

FIGURE 6

```
 83      ------------------------------------------------------------
 89      ANSMLSTIQKTFVTSSVTNHHFSNASQSFRSPILGVNAKIGYQNYFNDFIGLAYYGIIKY
108      ANSMLSTIQKTFVTSSVTNHHFSNASQSFRSPILGVNAKIGYQNYFNDFIGLAYYGIIKY
118      ANSMLSTIQKTFVTSSVTNHHFSNASQSFRSPILGVNAKIGYQNYFNDFIGLAYYGIIKY
         ************************************************************

 83      ------------------------------------------------------------
 89      NYAKAVNQKVQQLSYGGGIDLLLDFITTYSNKNSPTGIQTKRNFSSSFGIFGGLRGLYNS
108      NYAKAVNQKVQQLSYGGGIDLLLDFITTYSNKNSPTGIQTKRNFSSSFGIFGGLRGLYNS
118      NYAKAVNQKVQQLSYGGGIDLLLDFITTYSNKNSPTGIQTKRNFSSSFGIFGGLRGLYNS
         ************************************************************

 83      ------------------------------------------------------------
 89      YYVLNKVKGSGNLDVATGLNYRYKHSKYSVGISIPLIQRKASVVSSGGDYTNSFVFNEGA
108      YYVLNKVKGSGNLDVATGLNYRYKHSKYSVGISIPLIQRKASVVSSGGDYTNSFVFNEGA
118      YYVLNKVKGSGNLDVATGLNYRYKHSKYSVGISIPLIQRKASVVSSGGDYTNSFVFNEGA
         ************************************************************

 83      ------------
 89      SHFKVFFNYGGCF
108      SHFKVFFNYGWVF
118      SHFKVFFNYGWVF
         ************
```

FIGURE 6 (Cont'd)

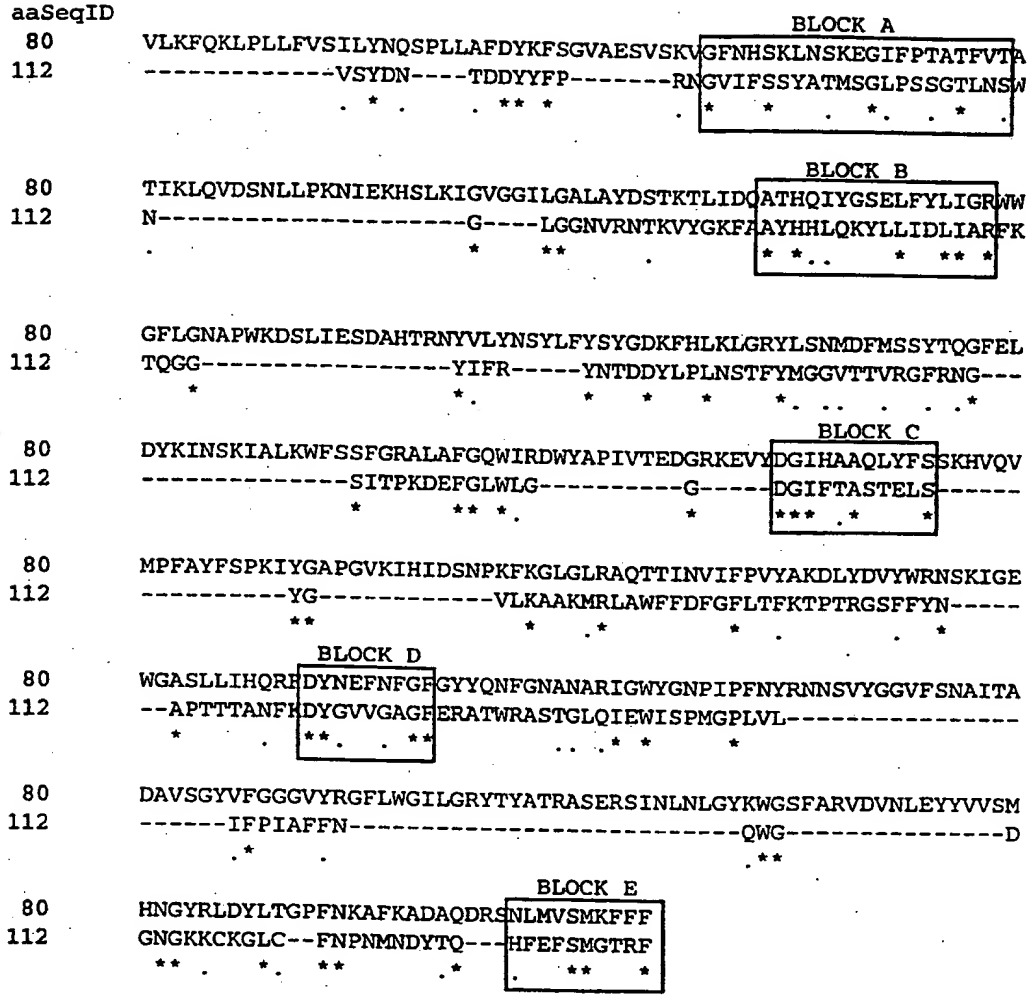
aaSeqID
BLOCK A
80 VLKFQKLPLLFVSILYNQSPLLAFDYKFSGVAESVSKVGFNHSKLNSKEGIFPTATFVTA
112 ----------VSYDN----TDDYYFP-------RNGVIFSSYATMSGLPSSGTLNSW
BLOCK B
80 TIKLQVDSNLLPKNIEKHSLKIGVGGILGALAYDSTKTLIDQATHQIYGSELFYLIGRWW
112 N----------G----LGGNVRNTKVYGKFAAYHHLQKYLLIDLIARFK
80 GFLGNAPWKDSLIESDAHTRNYVLYNSYLFYSYGDKFHLKLGRYLSNMDFMSSYTQGFEL
112 TQGG----------YIFR-----YNTDDYLPLNSTFYMGGVTTVRGFRNG---
BLOCK C
80 DYKINSKIALKWFSSFGRALAFGQWIRDWYAPIVTEDGRKEVYDGIHAAQLYFSSKHVQV
112 ----------SITPKDEFGLWLG----------G-----DGIFTASTELS------
80 MPFAYFSPKIYGAPGVKIHIDSNPKFKGLGLRAQTTINVIFPVYAKDLYDVYWRNSKIGE
112 ----------YG----------VLKAAKMRLAWFFDFGFLTFKTPTRGSFFYN-----
BLOCK D
80 WGASLLIHQRFDYNEFNFGFGYYQNFGNANARIGWYGNPIPFNYRNNSVYGGVFSNAITA
112 --APTTTANFKDYGVVGAGFERATWRASTGLQIEWISPMGPLVL----------
80 DAVSGYVFGGGVYRGFLWGILGRYTYATRASERSINLNLGYKWGSFARVDVNLEYYVVSM
112 ------IFPIAFFN----------QWG----------D
BLOCK E
80 HNGYRLDYLTGPFNKAFKADAQDRSNLMVSMKFFF
112 GNGKKCKGLC--FNPNMNDYTQ---HFEFSMGTRF

FIGURE 7

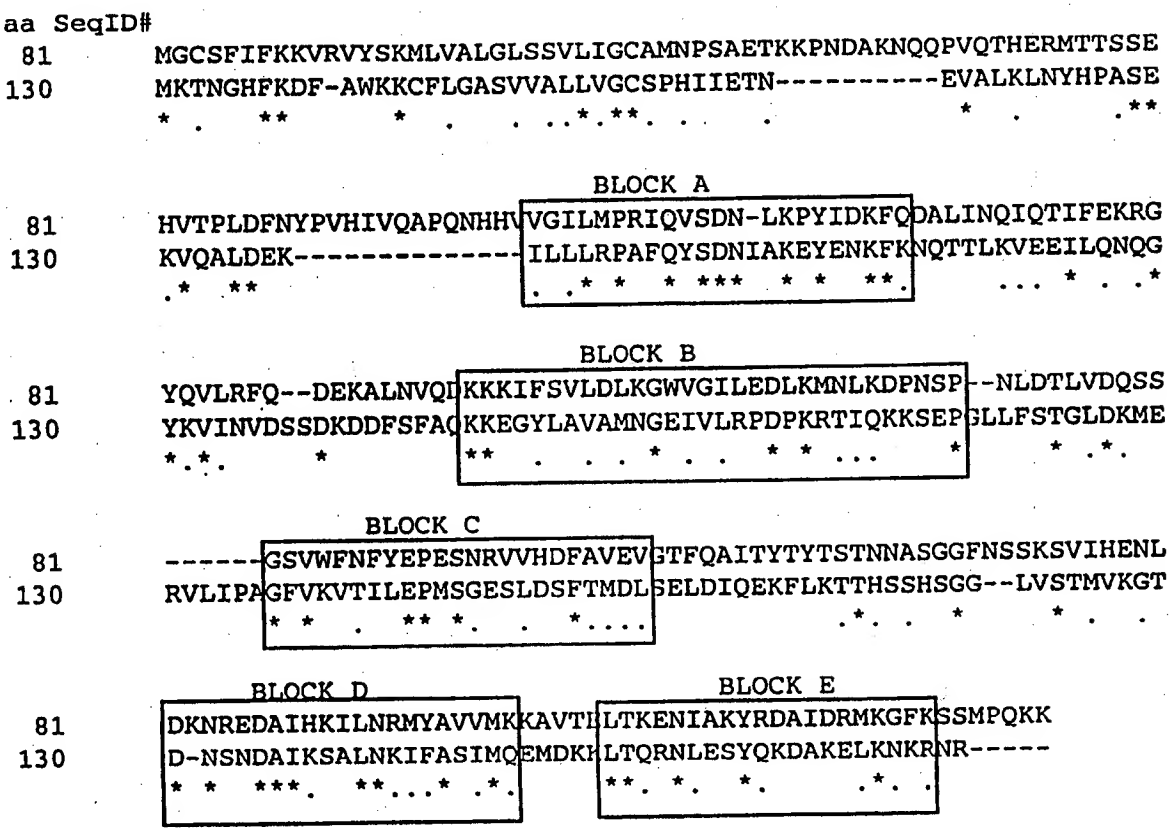
aa SeqID#
81 MGCSFIFKKVRVYSKMLVALGLSSVLIGCAMNPSAETKKPNDAKNQQPVQTHERMTTSSE
130 MKTNGHFKDF-AWKKCFLGASVVALLVGCSPHIIETN----------EVALKLNYHPASE
BLOCK A
81 HVTPLDFNYPVHIVQAPQNHHVVGILMPRIQVSDN-LKPYIDKFQDALINQIQTIFEKRG
130 KVQALDEK----------ILLLRPAFQYSDNIAKEYENKFKNQTTLKVEEILQNQG
BLOCK B
81 YQVLRFQ--DEKALNVQDKKKIFSVLDLKGWVGILEDLKMNLKDPNSP--NLDTLVDQSS
130 YKVINVDSSDKDDFSFAQKKEGYLAVAMNGEIVLRPDPKRTIQKKSEPGLLFSTGLDKME
BLOCK C
81 ------GSVWFNFYEPESNRVVHDFAVEVGTFQAITYTYTSTNNASGGFNSSKSVIHENL
130 RVLIPAGFVKVTILEPMSGESLDSFTMDLSELDIQEKFLKTTHSSHSGG--LVSTMVKGT
BLOCK D
81 DKNREDAIHKILNRMYAVVMKKAVTELTKENIAKYRDAIDRMKGFKSSMPQKK
130 D-NSNDAIKSALNKIFASIMQEMDKHLTQRNLESYQKDAKELKNKRNR-----
BLOCK E

FIGURE 8

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/19575

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :A01N 43/04; A61K 31/70; C12Q 1/68
US CL :514/44; 435/6
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 514/44; 435/6

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
GENEBANK

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
NONE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | TAYLOR, et al. Construction of a *Helicobacter pylori* Genome Map and Demonstration of Diversity at the Genome Level. Journal of Bacteriology. November 1992, Vol. 174, No. 21, pages 6800-6806, see entire document. | 1-65 |
| A | AKOPYANZ, et al. DNA diversity among clinical isolates of *Helicobacter pylori* detected by PCR-based RAPD fingerprinting. Nucleic Acids Research. 1992, Vol. 20, No. 19, pages 5137-5142, see entire document. | 1-65 |

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

| | | | |
|---|---|---|---|
| * | Special categories of cited documents: | "T" | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention |
| "A" | document defining the general state of the art which is not considered to be of particular relevance | "X" | document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone |
| "E" | earlier document published on or after the international filing date | "Y" | document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "L" | document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | "&" | document member of the same patent family |
| "O" | document referring to an oral disclosure, use, exhibition or other means | | |
| "P" | document published prior to the international filing date but later than the priority date claimed | | |

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 27 FEBRUARY 1998 | 13 MAR 1998 |
| Name and mailing address of the ISA/US<br>Commissioner of Patents and Trademarks<br>Box PCT<br>Washington, D.C. 20231<br>Facsimile No. (703) 305-3230 | Authorized officer<br>GINNY PORTNER<br>Telephone No. (703) 308-0196 |

Form PCT/ISA/210 (second sheet)(July 1992)*

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/19575

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
1-65, SEQ. ID Nos. 1, 7, 8, 11, 37, 39, 43, 45, 55, 61, 74, 80, 81 and 112

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/19575

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-26, 47, 49, 51, 53, 55, 57, 59, and 61, drawn to no fewer than 135 nucleic acid molecules, vectors containing the nucleic acid molecules, DNA encoding fragments of the polypeptides encoded by the no fewer than 135 different DNAs, organism transformed with the nucleic acid molecules, vaccines and methods of producing polypeptides encoded by the no fewer than 135 different nucleic acid molecules.

Group II, claim(s) 27-46, 48, 50, 52, 54, 56, 58, 60, and 62-65 are, drawn to no fewer than 73 polypeptides encoded by a subset of the encoding DNA mentioned in Group I.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

Group I contains a separate DNA species for each sequence mentioned. Therefore, there is a minimum of 135 species.

Group II contains at least one polypeptide for each DNA sequence mentioned. Therefore, this is a minimum of 73 species in Group II.

For either Group that applicant elects, a total of 10 (ten) specified sequences will be searched and no more than 4 (four) specified sequences will be searched for each additional fee paid; if no additional fee is paid and no election indicated the first 10 sequences appearing in Group I will be searched.

and it considers that the International Application does not comply with the requirements of unity of invention (Rules 13.1, 13.2 and 13.3) for the reasons indicated below:

The inventions listed as Groups I and II do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The polypeptide encoding DNAs, vectors containing them, organisms transformed with them and methods of polypeptide production using them of Group I are materially different from each other and are therefore independent from the polypeptides of Group II. Additionally, none of the products or methods of Group I is needed to make the polypeptides of Group II.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: There is no relationship between or among the various nucleotide and amino acid sequences mentioned in the claims.